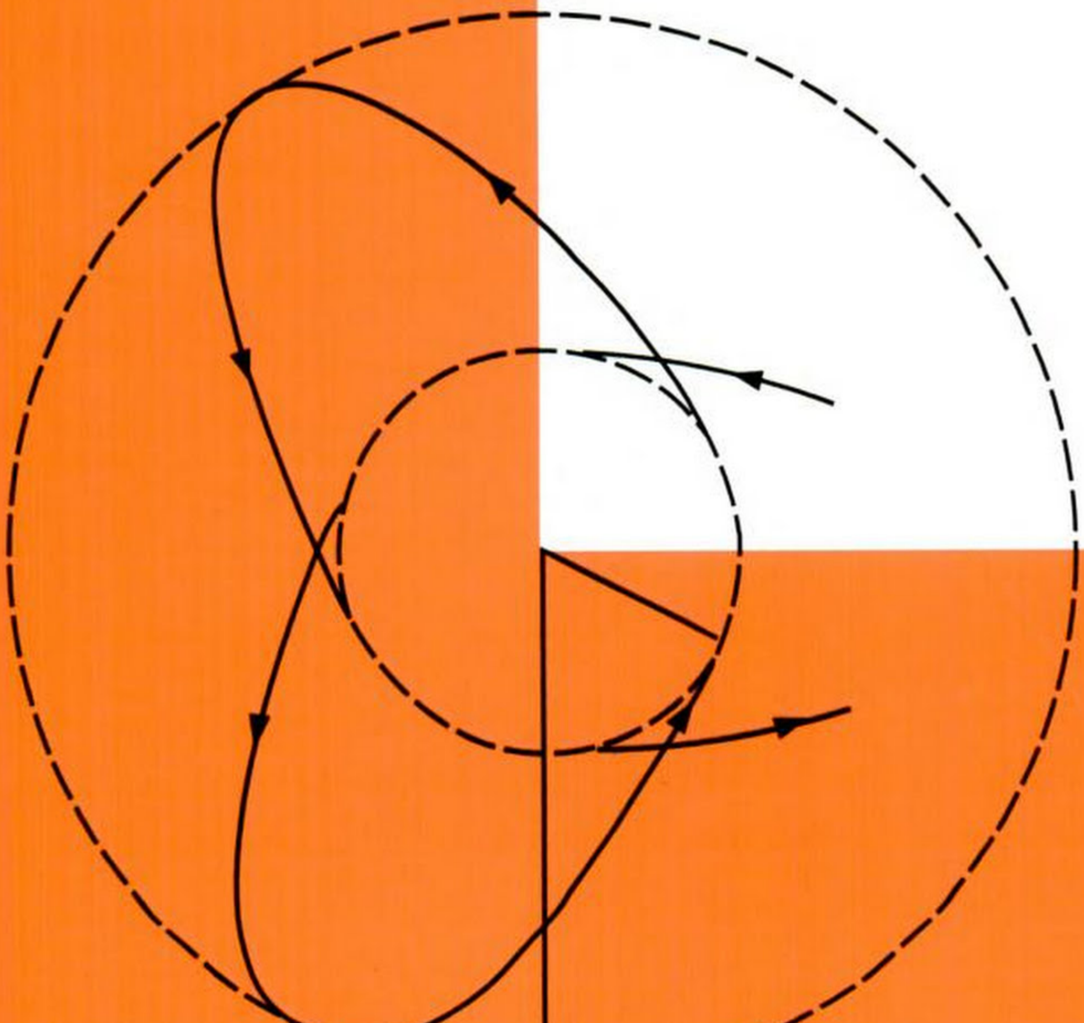
INTRODUCTION TO CLASSICAL MECHANICS

R G Takwale & P S Puranik



INTRODUCTION TO CLASSICAL MECHANICS

R G TAKWALE
P S PURANIK

Tata McGraw-Hill Publishing Company Limited
New Delhi

McGraw-Hill Offices

New Delhi New York St Louis San Francisco Auckland Bogotá Caracas
Lisbon London Madrid Mexico City Milan Montreal
San Juan Singapore Sydney Tokyo Toronto

Tata McGraw-Hill

34th reprint 2006
RXLLCRBKRZXQY

ISBN 0-07-096617-6

Published by Tata McGraw-Hill Publishing Company Limited, 7 West Patel Nagar, New Delhi 110 008, and printed at Ram Book Binding House, New Delhi 110 020

The McGraw-Hill Companies

Contents

INTRODUCTION TO
CLASSICAL MECHANICS

1 Vector Algebra

In physics, we try to find relationships between various physical quantities whose values may be determined experimentally.

Many quantities in physics can be completely specified by giving their magnitude alone. Such quantities are mass, density, temperature, etc. These are called *scalar quantities*. Many other quantities however, require, in addition to their magnitude, direction for their complete specification. These are called *vector quantities*. Displacement, force, electric field intensity, etc., are examples of vector quantities.

Many equations in physics assume a compact form when written in vector notation. The relationship between various quantities involved in equation is revealed immediately when these are written in vector form. Historically, vector notation became widely used with the advent of Maxwell's electromagnetic theory in which the above advantages are clearly seen.

Sometimes, we come across physical phenomenon in which we can associate a particular value of the variable with each point in a given region of the space. Such a region of the space is called a *field*. If the variable describes a scalar quantity, the field is called a *scalar field*. For example a temperature field around a hot body. If the variable describes a vector quantity, the field is called a *vector field*, for example a magnetic field.

A vector quantity may be geometrically represented by a straight line (i) having a length proportional to the magnitude of the vector quantity, and (ii) drawn in the same direction and sense as that of the given vector quantity.

In this book, we shall use the following notation for the representation of the vector quantities: (i) Bold-faced letters are used to represent the vector quantities. Thus, **A** represents 'vector *A*'. Similarly, **PQ** means a vector represented by a segment *PQ* of a straight line directed from *P* to *Q*. (ii) While writing the magnitude of vector quantities, italic letters are used. Thus, *A* represents the magnitude of vector **A**.

The direction of a vector is shown by an arrow and is reversed by reversing the direction of the arrow-head (Fig. 1.1). Thus **PQ** represents vector **A**, while **QP** represents vector **−A**. Referring to Fig. 1.2 one observes that **PQ**, $\mathbf{P_1Q_1}$ and $\mathbf{P_2Q_2}$ represent the same vector **A**. But, such a parallel or transverse translation of a vector may not always be physically equivalent. In order to consider this point, we have to realise that the vector quantities occurring in nature can be described by *free* or *sliding* or *bound* vectors. A free vector can be slid along the line of action or can be shifted parallel to itself without affecting the physical situation. A force acting on a rigid body, if slid along its line of action, will not change the rotation. However, if the force is shifted parallel to itself, it would produce a different torque and hence a different rotation. An example of a bound vector is an intensity vector of a field due to a point charge or mass. Intensity is a function of coordinates and the intensity vector cannot be shifted or slid since it will amount to change of coordinates and in turn a change in intensity.

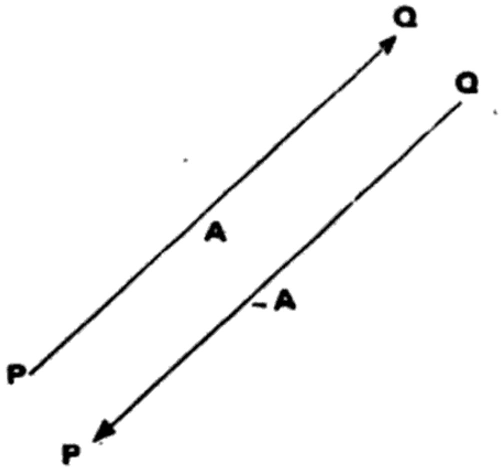


Fig. 1.1 Vector A and its negative vector

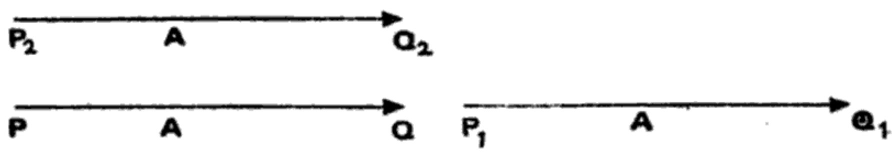


Fig. 1.2 Parallel or transverse translation of a vector

While adding, subtracting or multiplying in vector algebra, we implicitly assume that the physical situation from which the vectors are taken is unchanged.

1.1 ADDITION OF VECTORS

Two vectors **A** and **B** can be added by using the triangle law or the

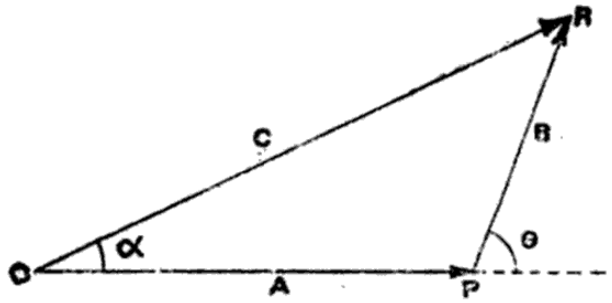


Fig. 1.3 Triangle law of addition of two vectors: **C** = **A** + **B**

parallelogram law of vectors (Figs. 1.3 and 1.4). The resultant **C** can be

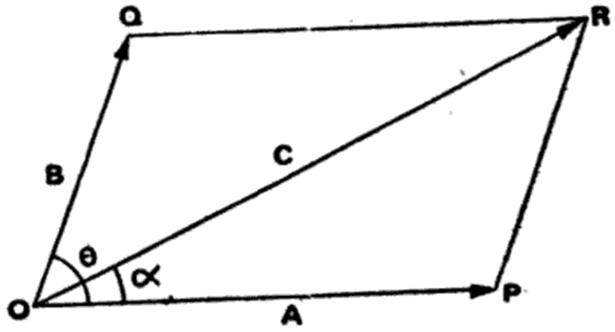


Fig. 1.4 Parallelogram law of addition of two vectors: **C** = **A** + **B**

shown to possess a magnitude

$$C = \sqrt{A^2 + B^2 + 2AB\cos\theta} \tag{1.1}$$

and its direction is exprcssed in terms of angle α made by **C** with **A** and given by

$$\tan\alpha = \frac{B\sin\theta}{A + B\cos\theta} \tag{1.2}$$

If more than two vectors are to be added we use a polygon law which is a mere repetition of the triangle law.

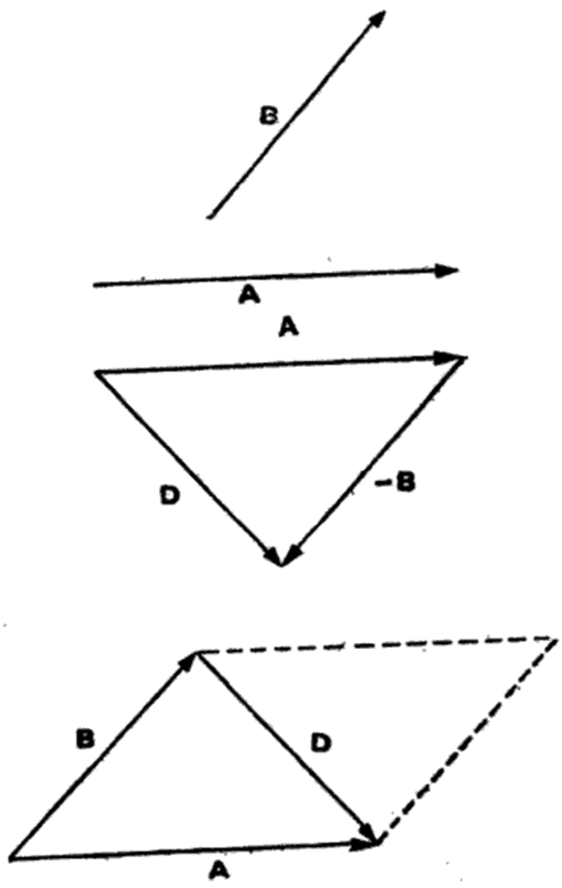


Fig. 1.5 Subtraction of vectors: **D** = **A** − **B**

The subtraction of vector **B** from **A** is carried out according to the equation

$$\mathbf{D} = \mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B}) \tag{1.3}$$

Thus, if **A** is added to $(-\mathbf{B})$, we get $\mathbf{D} = \mathbf{A} - \mathbf{B}$ (Fig. 1.5).

The addition of vectors is commutative, i.e.

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \tag{1.4}$$

Similarly, addition of vectors is associative, i.e.

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C} \tag{1.5}$$

1.2 EQUALITY OF VECTORS

The vector equation $\mathbf{A} = \mathbf{B}$ indicates that the magnitudes as well as directions of **A** and **B** are identical. Then,

$$\mathbf{A} - \mathbf{B} = \mathbf{0} \tag{1.6}$$

Since this is a vector equation, the right-hand side must also be a vector and is called a *null or zero-vector*. A null vector has a zero magnitude.

1.3 UNIT VECTOR

Vector **A** can be written as

$$\mathbf{A} = \hat{\mathbf{e}}_A A \tag{1.7}$$

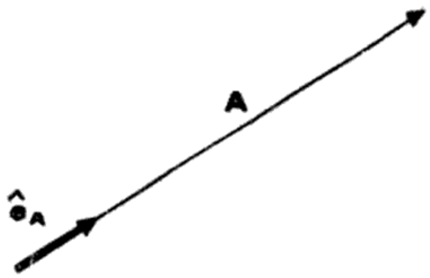


Fig. 1.6 Unit vector $\hat{\mathbf{e}}_A$ gives $\mathbf{A} = \hat{\mathbf{e}}_A A$

where A is the magnitude of **A** and $\hat{\mathbf{e}}_A$ is a vector having a unit magnitude and is drawn in the same direction as that of **A**. It is called a unit vector in the direction of a given vector (Fig. 1.6). The representation of a vector in terms of a unit vector is very useful in vector algebra. For example, if a vector **A** is multiplied by a pure number n, we get

$$n\mathbf{A} = \hat{\mathbf{e}}_A(nA) \tag{1.8}$$

The result represents a vector of magnitude which is n times greater than that of the original vector, but its direction remains unaffected.

1.4 PRODUCT OF TWO VECTORS

Consider force **F** which produces displacement **r** in a direction that makes angle θ with **F** (Fig. 1.7). In this process, work is said to be done by the

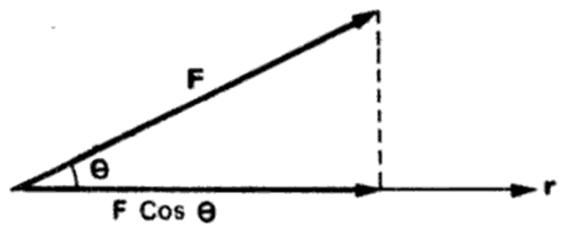


Fig. 1.7 Work done by force $W = \mathbf{F}\cdot\mathbf{r} = Fr\cos\theta$

force which is given by

$$W = Fr \cos \theta \tag{1.9}$$

Here vector quantities **F** and **r** are multiplied giving work W which is a scalar quantity.

Now consider the torque produced by force **F** about point O (Fig. 1.8).

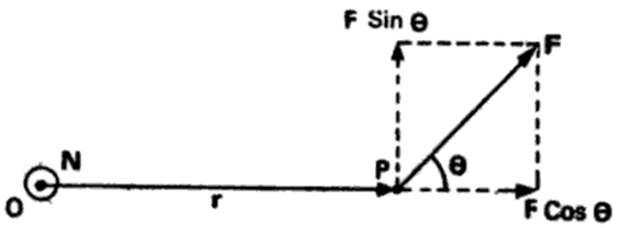


Fig. 1.8 Torque $\mathbf{N} = \mathbf{r} \times \mathbf{F}$

Let the force act at point P such that position vector $\mathbf{OP} = \mathbf{r}$ makes angle θ with the direction of **F**. Then, the magnitude of the torque is given by

$$N = F \sin \theta \, r \quad \text{or} \quad N = Fr \sin \theta \tag{1.10}$$

But the torque produces rotation which has rotational sense, clockwise or anticlockwise, and hence is regarded as a vector quantity. The torque vector is represented by means of a directed straight line perpendicular to the plane formed by **F** and **r**. In Fig. 1.8, torque vector **N** is at right angles to the plane of the figure and its direction is towards the reader. This convention is in accordance with the right-hand screw rule applied to the anticlockwise rotation that would be produced in the situation shown in Fig. 1.8. (This direction is represented by symbol $\odot$ at point O in Fig. 1.8. The opposite direction of rotation will be shown by $\otimes$ at the point on the axis of rotation.)

These two illustrations show that the product of two vector quantities is either a scalar quantity or a vector quantity. Accordingly, we define: (a) a scalar product or dot product, and (b) a vector product or cross product of two vectors.

(a) Scalar or Dot Product of Two Vectors

The scalar or dot product of two vectors **A** and **B** is defined as

$$\mathbf{A} \cdot \mathbf{B} = AB \cos \theta \tag{1.11}$$

where A and B are the scalar magnitudes of **A** and **B** and θ is the angle between the two vectors.

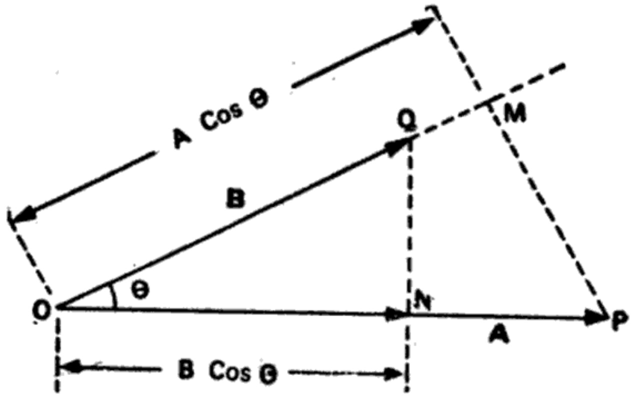
It is clear from Fig. 1.9, that $B \cos \theta$ is the projection of vector **B** along the direction of vector **A** or $A \cos \theta$ is the projection of vector **A** in the direction of vector **B**. Hence, we can write

$$\mathbf{A} \cdot \mathbf{B} = AB \cos \theta = A \cos \theta \, B = \mathbf{B} \cdot \mathbf{A} \tag{1.12}$$

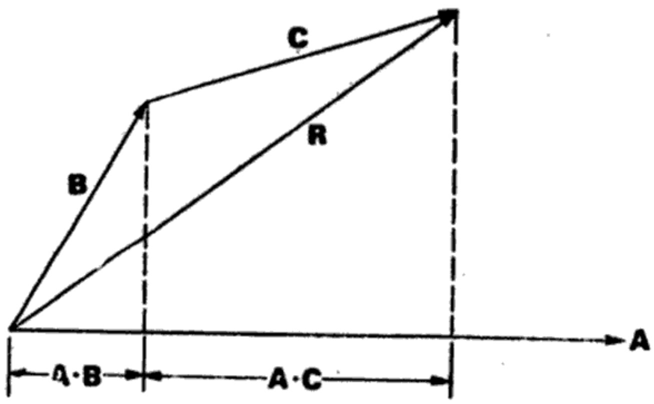
Thus, the dot product of two vectors is commutative. This is due to the fact that $\cos(-\theta) = \cos \theta$.

Here we use the usual sign convention for the angles. Thus the angle

described in an anticlockwise direction is positive, while that described in a clockwise direction is negative.



(a)



(b)

Fig. 1.9 (a) Scalar or dot product of two vectors; (b) geometrical proof of the law of distribution for dot product: $\mathbf{A}\cdot(\mathbf{B}+\mathbf{C}) = \mathbf{A}\cdot\mathbf{B} + \mathbf{A}\cdot\mathbf{C}$

In general, we can interpret the dot product of two vectors as the product of the scalar magnitude of one vector and the projection of the other vector in the direction of the first vector.

If the angle between the two vectors is $\frac{\pi}{2}$, i.e. the two vectors **A** and **B** are perpendicular to each other, then

$$\mathbf{A}\cdot\mathbf{B} = AB\cos\frac{\pi}{2} = 0 \tag{1.13}$$

In this case, vectors **A** and **B** are said to be *orthogonal*.

If vectors **A** and **B** are parallel ($\theta = 0$) or antiparallel ($\theta = \pi$), then

$$\mathbf{A}\cdot\mathbf{B} = AB \quad \text{or} \quad \mathbf{A}\cdot\mathbf{B} = -AB \tag{1.14}$$

respectively.

From the definition of the dot product of two vectors, we can write $\mathbf{A}\cdot\mathbf{A} = AA\cos 0 = A^2$. Hence

$$A = \sqrt{\mathbf{A}\cdot\mathbf{A}} = \sqrt{A^2} \tag{1.15}$$

The definition of the dot product given in equation (1.11) can also be used to prove that the dot product is distributive (Fig. 1.9b). Thus

$$\mathbf{A}\cdot(\mathbf{B}+\mathbf{C}) = \mathbf{A}\cdot\mathbf{B} + \mathbf{A}\cdot\mathbf{C} \tag{1.16}$$

Several illustrations of the dot product of two vectors can be mentioned. For example: work done $W = \mathbf{F}\cdot\mathbf{r}$, electric potential energy $U = q\int_0^l \mathbf{E}\cdot d\mathbf{l}$, electric flux $\Phi_E = \int_\sigma \mathbf{E}\cdot d\boldsymbol{\sigma}$, where $\mathbf{E}$ is the electric field intensity, $d\mathbf{l}$ is the line element and $d\boldsymbol{\sigma}$ is the surface element.

(b) Vector or Cross Product of Two Vectors

The vector or the cross product of the two vectors **A** and **B** is vector **C** in a direction perpendicular to the plane formed by **A** and **B** and has magnitude $AB \sin\theta$.

Thus,

$$\mathbf{C} = \mathbf{A}\times\mathbf{B} = \hat{\mathbf{e}}_c\, AB \sin\theta \tag{1.17}$$

where $A \equiv |\mathbf{A}|$, $B \equiv |\mathbf{B}|$, and θ is the angle between the directions of **A** and **B**. The sense of **C** is fixed by the sense in which the tip of a right-hand screw would advance if the head of such a screw is rotated from the

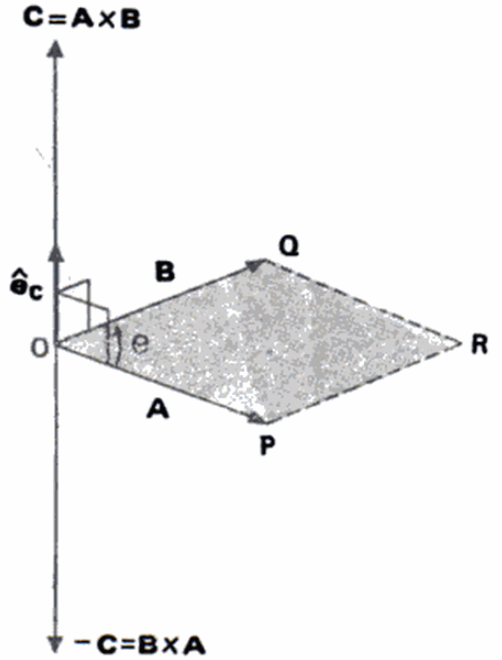


Fig. 1.10 The cross product of two vectors is not commutative

direction of **A** to the direction of **B**. Thus, vectors **A**, **B** and **C** form a right-handed triad. It is obvious from Fig. 1.10, that

$$\mathbf{A}\times\mathbf{B} = -\mathbf{B}\times\mathbf{A} \tag{1.18}$$

i.e., the cross product of two vectors is non-commutative.

Combining equations (1.17) and (1.18) we can write

$$\mathbf{C} = \mathbf{A}\times\mathbf{B} = -\,\mathbf{B}\times\mathbf{A} = \hat{\mathbf{e}}_c AB \sin\theta \tag{1.19}$$

where $\hat{\mathbf{e}}_c$ is a unit vector in the direction of **C** and is perpendicular to the

plane formed by **A** and **B**. Magnitude of the vector product of **A** and **B**, viz. $AB \sin \theta$, is the area of the parallelogram *OPRQ* formed with sides **A** and **B**. This fact is, therefore, used to represent area as a vector quantity.

If the two vectors **A** and **B** are perpendicular to each other,

$$\mathbf{C} = \mathbf{A} \times \mathbf{B} = \hat{\mathbf{e}}_c\, AB \sin \frac{\pi}{2} = \hat{\mathbf{e}}_c\, AB \tag{1.20}$$

If the two vectors **A** and **B** are parallel ($\theta = 0$) or antiparallel ($\theta = \pi$),

$$\left.\begin{aligned} \mathbf{C} &= \mathbf{A} \times \mathbf{B} = \hat{\mathbf{e}}_c AB \sin \theta = 0 \\ \text{or} \quad \mathbf{C} &= \mathbf{A} \times \mathbf{B} = \hat{\mathbf{e}}_c AB \sin \pi = 0 \end{aligned}\right\} \tag{1.21}$$

It is also obvious that $\mathbf{A} \times \mathbf{A} = 0$.

As shown in article 1.8 the vector product is distributive, i.e. if **A**, **B** and **C** are the given vectors,

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C} \tag{1.22}$$

Several illustrations of the cross product of two vectors can be mentioned. For example, torque $\mathbf{N} = \mathbf{r} \times \mathbf{F}$, angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, magnetic force on a moving charge q_0 is $\mathbf{F}_m = q_0\mathbf{v} \times \mathbf{B}, \ldots$, etc. Here, **F** is the force, **p** the momentum and **B** the magnetic induction.

1.5 RESOLUTION OF A VECTOR

We have already seen that two or more vectors can be added to give a single resultant vector. This process is often called the *composition* of vectors. In the reverse process, called the *resolution* of a vector, we find two or more vectors which together would produce the same effect as that produced by the given vector. These vectors are then called the components of the given vector.

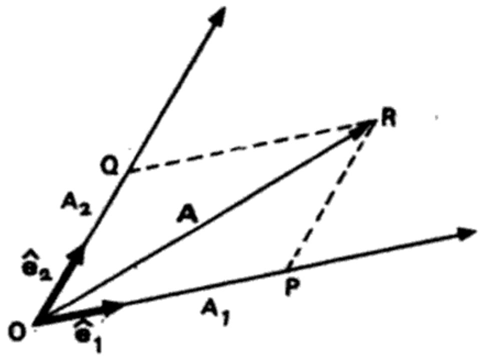


Fig. 1.11 Resolution of a vector into two components A_1 and A_2 in a plane containing A

Consider vector **A** (Fig. 1.11). Let us resolve it in a plane containing **A** along any two directions along which we take unit vectors $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$. For this we form the parallelogram OPRQ in Fig. 1.11 with vector **A** as the diagonal. Hence vector **A** can be considered the combination of the two vectors $\mathbf{OP} = \hat{\mathbf{e}}_1 A_1$ and $\mathbf{OQ} = \hat{\mathbf{e}}_2 A_2$. Thus, **OP** and **OQ** are the components of vector **A** along the directions of unit vectors $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ respectively.

For the magnitude of vector **A** we can write

$$A^2 = A_1^2 + A_2^2 + 2A_1A_2 \cos \angle POQ \tag{1.23}$$

Thus, any vector **A** can be expressed in terms of two components along any two reference lines in the plane containing **A**. Unit vectors $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ serve as reference or coordinate axes in two dimensions. The process described above can be generalised to three dimensions by introducing the third unit vector $\hat{\mathbf{e}}_3$ which will obviously not be in the plane of $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$.

The coordinate axes just mentioned are called the oblique coordinate axes. These are found to be rather inconvenient. In a special case, the unit vectors $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ are taken to be perpendicular to each other. Then, we have

$$A^2 = A_1^2 + A_2^2 \tag{1.24}$$

In this case, the coordinate system is said to be orthogonal. The cartesian coordinate system (Fig. 1.12) of mutually perpendicular axes is an orthogonal coordinate system.

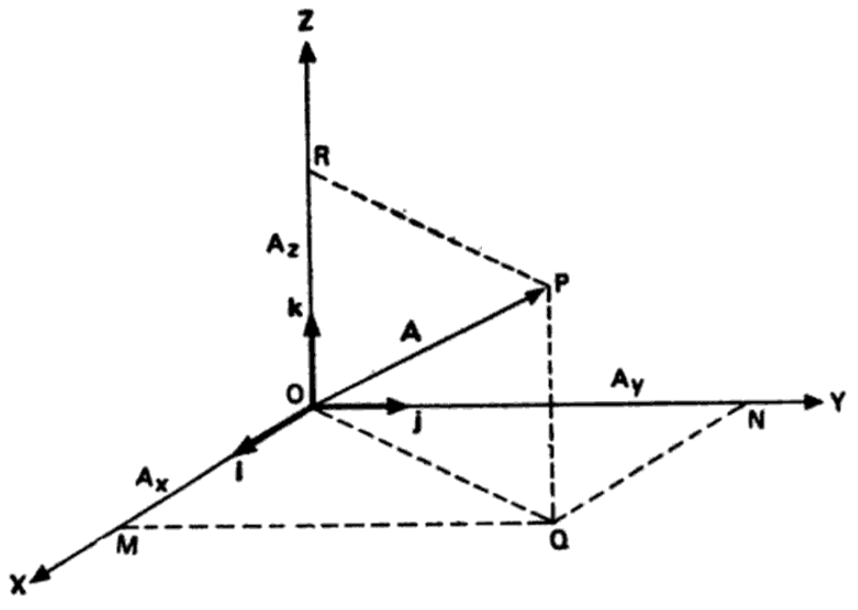


Fig. 1.12 Unit vectors **i**, **j**, **k** in a right-handed cartesian coordinate system

We follow the usual practice of denoting the unit vectors along the x, y and z-axes by **i**, **j** and **k** respectively. Then, vector **A** can be written as

$$\mathbf{A} = \mathbf{i}A_x + \mathbf{j}A_y + \mathbf{k}A_z \tag{1.25}$$

where A_x, A_y and A_z are the projections of vector **A** along the x, y and z-axes respectively.

Any vector in the three-dimensional space can be expressed in terms of unit vectors **i**, **j** and **k**. Hence, unit vectors **i**, **j** and **k** are said to *span* the whole space. Secondly, unit vectors **i**, **j** and **k** are said to be *linearly independent*. This is because any vector equation of the form

$$\mathbf{A} = \mathbf{i}A_1 + \mathbf{j}A_2 + \mathbf{k}A_3 = 0$$

will mean $A_1 = 0$, $A_2 = 0$ and $A_3 = 0$. For a given set of unit vectors **i**, **j**, **k**, we will not be able to express any vector, say $\mathbf{i}A_1$, along **i** as a sum of two vectors along the directions of **j** and **k**. Unit vectors **i**, **j** and **k** are said to form the *basis* of the three-dimensional space.

From Fig. 1.12, we can write in an orthogonal coordinate system

$$A = \sqrt{A_x^2 + A_y^2 + A_z^2} \tag{1.26}$$

which is simply the three-dimensional generalisation of equation (1.24).

Quantities $\frac{A_x}{A}$, $\frac{A_y}{A}$ and $\frac{A_z}{A}$ are the direction cosines l, m and n of vector **A** and $l^2 + m^2 + n^2 = 1$, again gives equation (1.26).

We shall always use a *right-handed coordinate system*. In this system, if the x-axis is turned towards the y-axis through the smaller angle $\left(= \frac{\pi}{2}\right)$ between them, the tip of the right-handed screw will advance along the z-axis.

If, however, the x-axis is turned towards the y-axis through the smaller angle $\left(= \frac{\pi}{2}\right)$ between them and if the tip of the *left-handed screw*

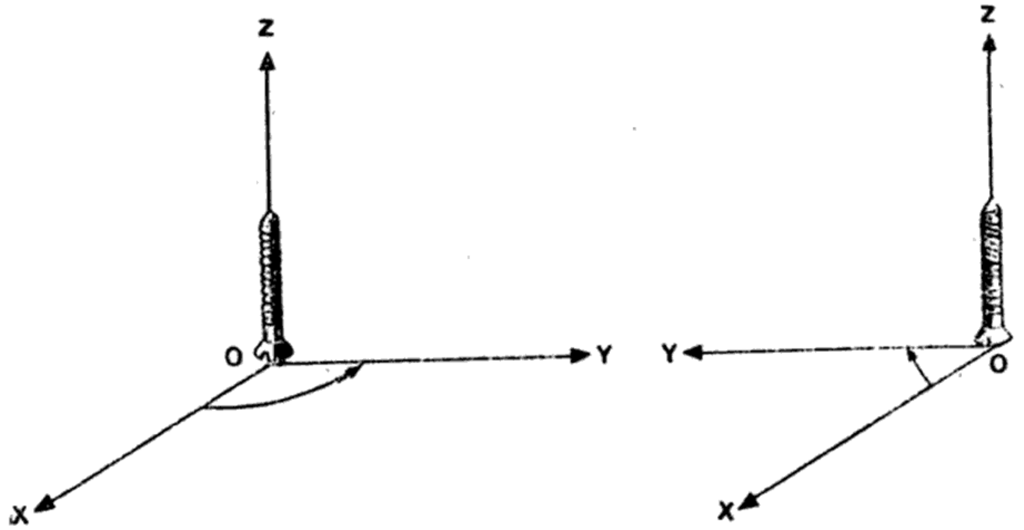


Fig. 1.13 (a) Right-handed coordinate system; (b) left-handed coordinate system.

advances in the direction of the z-axis, the system is called the *left-handed coordinate system*. These systems are shown in Fig. 1.13. It should be noted that these two systems cannot be made coincident by simply rotating them. One of the axes, i.e., the y-axis has an opposite direction and is said to be reflected in analogy with the similar effect observed in mirror reflection.

1.6 DEFINITION OF A VECTOR IN TERMS OF ITS COMPONENTS

We have defined a vector quantity as a quantity which requires magni-

tude (with a suitable unit) and direction for its complete specification. We can also define it in terms of its three components because the three components of a vector in a given three-dimensional coordinate system determine uniquely the magnitude and direction of the given vector. The components are arranged in the order **i**, **j** and **k** (or $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$) and are called *base vectors*. Thus, vector **A** is given by a set of three ordered numbers A_x, A_y and A_z and is written as

$$\mathbf{A} = (A_x, A_y, A_z) \tag{1.27}$$

This method of defining a vector has an obvious advantage in that we can generalise the concept of a vector to spaces having dimensions more than three. Thus, in a four-dimensional space,

$$\mathbf{A} = (A_1, A_2, A_3, A_4) \tag{1.28}$$

where suffixes 1, 2, 3 and 4 are used to indicate the orthogonal coordinate axes. Such a four-dimensional space is considered in the Special Theory of Relativity wherein time is taken as the fourth coordinate (taken imaginary in order to retain orthogonality) in addition to spatial coordinates x, y and z.

The generalisation to n dimensions or infinite dimensions becomes necessary in many problems such as representation of wave vectors in quantum mechanics. Another advantage of such a representation is that a vector with n components can be represented by a column matrix and the matrix theory can be applied freely to express physical equations.

In dealing with an n dimensional space, the symbol $\hat{\mathbf{e}}_i$, where $i = 1, 2, 3, \ldots, n$, is used for the base vectors. Then, vector **A** is written as

$$\mathbf{A} = (A_1, A_2, A_3, \ldots, A_n) \tag{1.29}$$

and the components, in general, may be complex.

1.7 VECTOR ALGEBRA IN TERMS OF THE COMPONENTS

1. The addition or subtraction of two vectors **A** and **B** can now be written as

$$\mathbf{A} \pm \mathbf{B} = \mathbf{i}(A_x \pm B_x) + \mathbf{j}(A_y \pm B_y) + \mathbf{k}(A_z \pm B_z) \tag{1.30}$$

The result of equation (1.30) can be easily generalised for any number of vectors.

2. The scalar or dot product of two vectors can be written as

$$\mathbf{A}\cdot\mathbf{B} = A_xB_x + A_yB_y + A_zB_z \tag{1.31}$$

This is because

$$\mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} = 1, \tag{1.32}$$

and

$$\mathbf{i}\cdot\mathbf{j} = \mathbf{j}\cdot\mathbf{k} = \mathbf{k}\cdot\mathbf{i} = \ldots \text{etc.} = 0 \tag{1.33}$$

3. The cross product of two vectors **A** and **B** is written as

$$\mathbf{C} = \mathbf{A}\times\mathbf{B} = (\mathbf{i}A_x + \mathbf{j}A_y + \mathbf{k}A_z)\times(\mathbf{i}B_x + \mathbf{j}B_y + \mathbf{k}B_z)$$
$$= \mathbf{i}(A_yB_z - A_zB_y) + \mathbf{j}(A_zB_x - A_xB_z) + \mathbf{k}(A_xB_y - A_yB_x)$$

$$\text{i.e.} \qquad \mathbf{A} \times \mathbf{B} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \tag{1.34}$$

This is because

$$\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = 0 \tag{1.35}$$

and

$$\left.\begin{aligned} \mathbf{i} \times \mathbf{j} &= \mathbf{k}, & \mathbf{j} \times \mathbf{i} &= -\mathbf{k} \\ \mathbf{j} \times \mathbf{k} &= \mathbf{i}, & \mathbf{k} \times \mathbf{j} &= -\mathbf{i} \\ \mathbf{k} \times \mathbf{i} &= \mathbf{j}, & \mathbf{i} \times \mathbf{k} &= -\mathbf{j} \end{aligned}\right\} \tag{1.36}$$

It will be noticed from above that

$$C_x = (A_y B_z - A_z B_y) \ldots, \text{ etc.} \tag{1.37}$$

4. If the two vectors $\mathbf{A}$ and $\mathbf{B}$ are given by

$$\mathbf{A} = (A_1, A_2, A_3, \ldots, A_n)$$

and

$$\mathbf{B} = (B_1, B_2, B_3, \ldots, B_n)$$

then the dot product of $\mathbf{A}$ and $\mathbf{B}$ is written as

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} &= A_1 B_1 + A_2 B_2 + \ldots + A_n B_n \\ &= \sum_{i=1}^{n} A_i B_i \end{aligned} \tag{1.38}$$

For equation (1.38) to be true, the base vectors which span the n-dimensional space must satisfy the relation,

$$\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij} \tag{1.39}$$

where δ_{ij} is the Kronecker delta symbol and is defined as

$$\left.\begin{aligned} \delta_{ij} &= 1, \quad \text{if} \quad i = j \\ \text{and} \qquad \delta_{ij} &= 0, \quad \text{if} \quad i \neq j \end{aligned}\right\} \tag{1.40}$$

The notation of equation (1.40) can be used in the case of three-dimensional space also.

5. The results of equation (1.36) can be written in a single equation if we use unit vectors $\hat{\mathbf{e}}_i$ $(i = 1, 2, 3)$. Thus,

$$\hat{\mathbf{e}}_i \times \hat{\mathbf{e}}_j = \mathcal{E}_{ijk} \hat{\mathbf{e}}_k \tag{1.41}$$

where $i, j, k = 1, 2, 3$. Symbol $\mathcal{E}_{ijk}$ is the permutation symbol or Levi Civita density. It has the meaning,

$$\left.\begin{aligned} &\mathcal{E}_{ijk} = 0, \text{ if any two of the indices } i, j, k \text{ are equal.} \\ &\mathcal{E}_{ijk} = 1, \text{ if } i, j, k \text{ form an even number of} \\ &\qquad\qquad \text{permutations of } 1, 2, 3. \\ \text{and} \quad &\mathcal{E}_{ijk} = -1, \text{ if } i, j, k \text{ form an odd number of} \\ &\qquad\qquad \text{permutations of } 1, 2, 3. \end{aligned}\right\} \tag{1.42}$$

Consider for example $\mathcal{E}_{321}$. It can be transformed into $\mathcal{E}_{123}$ by the

following permutations,

$$\mathcal{E}_{321} \to \mathcal{E}_{312} \to \mathcal{E}_{132} \to \mathcal{E}_{123}$$

Thus, the number of permutations necessary for this purpose is *three*, i.e., *odd*. Hence, $\mathcal{E}_{321} = -1$.

Similarly,

$$\mathcal{E}_{122} = \mathcal{E}_{113} = \mathcal{E}_{233} = 0, \text{ etc.}$$

$$\mathcal{E}_{123} = \mathcal{E}_{231} = \mathcal{E}_{312} = +1$$

and

$$\mathcal{E}_{132} = \mathcal{E}_{321} = \mathcal{E}_{213} = -1$$

The values $+1$ or -1 of the permutation symbol $\mathcal{E}$ can be judged from cyclic or anticyclic order of distinct indices (i, j, k) respectively.

In this notation of $\mathcal{E}$, the cross product of two vectors is expressed as

$$\mathbf{C} = \mathbf{A} \times \mathbf{B} = (\sum_i \hat{\mathbf{e}}_i A_i) \times (\sum \hat{\mathbf{e}}_j B_j)$$

$$= \sum_{ij} \hat{\mathbf{e}}_i \times \hat{\mathbf{e}}_j A_i B_j$$

or

$$\mathbf{C} = \mathbf{A} \times \mathbf{B} = \sum_{ijk} \mathcal{E}_{ijk} \hat{\mathbf{e}}_k A_i B_j \tag{1.43}$$

1.8 SURFACE AREA AS A VECTOR

The vector product of two vectors **A** and **B** is defined by the equation,

$$\mathbf{C} = \mathbf{A} \times \mathbf{B} = \hat{\mathbf{e}}_c \, AB \sin \theta$$

But, $AB \sin \theta$ represents the area of a parallelogram whose sides are **A** and **B** (Fig. 1.14). Hence, we can represent the area of the parallelogram

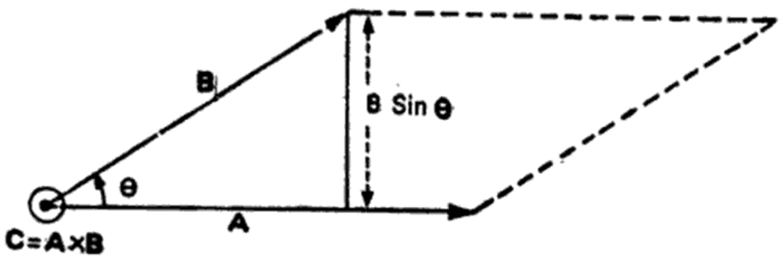


Fig. 1.14 Surface area as a vector

by vector **C**. When thus represented, the area of the parallelogram is shown by means of a vector drawn at right angles to the plane of the area. The direction of this area vector in Fig. 1.14 is towards the reader in accordance with the right-hand screw rule applied to the cross product $\mathbf{A} \times \mathbf{B}$.

But vector **C** does not indicate any shape of the area. Hence, any plane area can be represented by a straight line whose length is proportional to the magnitude of the area and is drawn perpendicular to the plane of the area. The arrow-head is pointed in the direction of the advance of the

tip of right-hanu screw whose head is rotated in the direction in which the bounding curve is traversed (Fig. 1.15a).

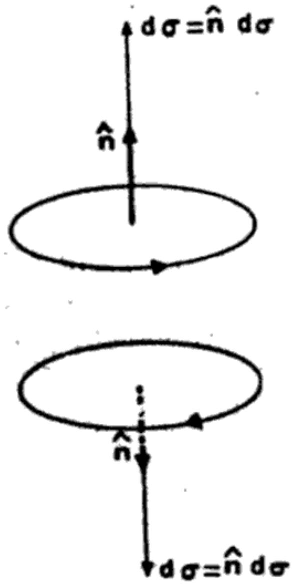


Fig. 1.15a Direc tion of area vector

If the surface under consideration is a part of a closed figure (say a tetrahedron), then the area vector of this surface is by convention represented by a straight line pointing outward, i.e., the boundary of each surface is supposed to be traversed in such a direction that, if the head of the right-hand screw is rotated in the same direction, its tip will advance outward. The vector representing the entire surface area of a closed surface is zero. This is because the projection of the entire surface area on any plane will give as much negative area vector as positive.

1.9 DISTRIBUTION LAW FOR VECTOR PRODUCT

To prove equation (1.22), let us consider the prism (Fig. 1.15b) with sides **A**, **B**, **A** + **B** and **C**. The outward vectors representing the areas of the faces of the prism are,

$$\text{Area } PQTS \Rightarrow \mathbf{C} \times \mathbf{A}, \qquad \text{Area } QRUT \Rightarrow \mathbf{C} \times \mathbf{B}$$

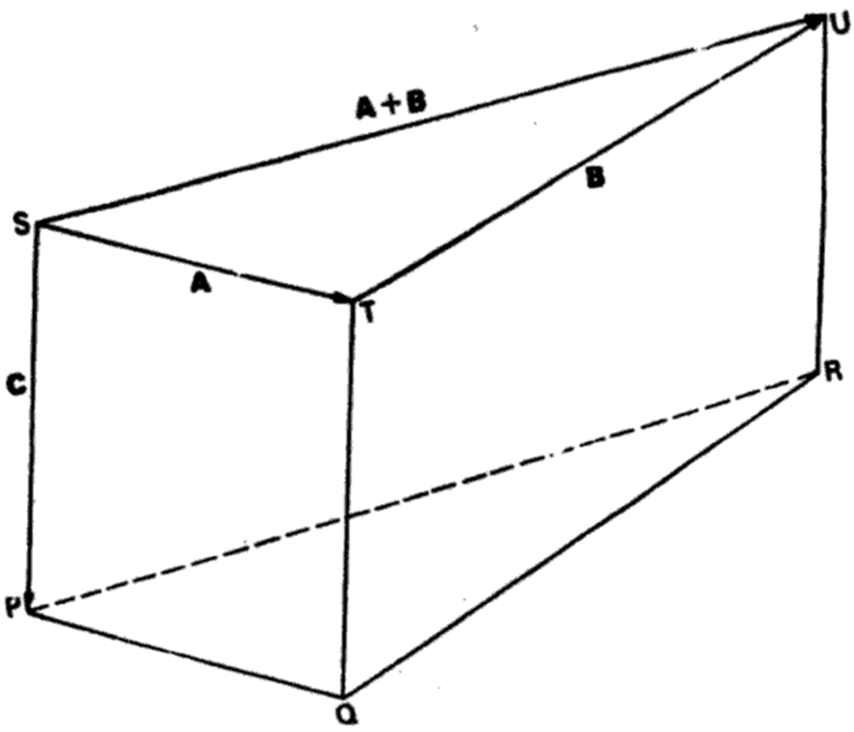


Fig. 1.15b Prism with sides A, B, A + B and C

Area $PRUS \Rightarrow (\mathbf{A} + \mathbf{B}) \times \mathbf{C}$, Area $STU \Rightarrow \frac{1}{2}\mathbf{A} \times \mathbf{B}$

and Area $PQR \Rightarrow \frac{1}{2}\mathbf{B} \times \mathbf{A}$

The vector sum of the entire surface area of the prism is zero. Hence,

$$\mathbf{C} \times \mathbf{A} + \mathbf{C} \times \mathbf{B} + (\mathbf{A} + \mathbf{B}) \times \mathbf{C} + \tfrac{1}{2}\mathbf{A} \times \mathbf{B} + \tfrac{1}{2}\mathbf{B} \times \mathbf{A} = 0$$

By using the anticommutative nature of the vector product given by equation (1.18), we get

$$\mathbf{C} \times \mathbf{A} + \mathbf{C} \times \mathbf{B} = -(\mathbf{A} + \mathbf{B}) \times \mathbf{C} = \mathbf{C} \times (\mathbf{A} + \mathbf{B})$$

which is equation (1.22).

1.10 DYADIC OR TENSOR OF RANK TWO

We know that whatever be the coordinate system chosen, one number is sufficient to describe a scalar completely while three numbers are required to describe a vector completely. From this point of view, it is convenient to regard the scalar as a tensor of rank zero (since it has in three-dimensional space $3^0 = 1$ component), while a vector is regarded as a tensor of rank one (since it has $3^1 = 3$ components). A quantity having $3^2 = 9$ components is called a tensor of rank two. Thus, in general, a tensor of rank n will have 3^n components.

In physics, we come across several quantities which are tensors of rank two. A tensor of rank two is called a *dyadic* or a *dyad* and can be expressed as a simple product of two vectors **A** and **B** without dot or cross. It has nine components given by

$$\begin{aligned}\mathbf{AB} = \mathbf{ii}A_xB_x + \mathbf{ij}A_xB_y + \mathbf{ik}A_xB_z + \mathbf{ji}A_yB_x + \mathbf{jj}A_yB_y + \mathbf{jk}A_yB_z \\ + \mathbf{ki}A_zB_x + \mathbf{kj}A_zB_y + \mathbf{kk}A_zB_z \qquad (1.44)\end{aligned}$$

Product **AB** is non-commutative, i.e.,

$$\mathbf{AB} \neq \mathbf{BA} \qquad (1.45)$$

A dyadic can be dotted with vector **C** either from the left or from the right. Thus,

$$\mathbf{C}\cdot(\mathbf{AB}) = (\mathbf{C}\cdot\mathbf{A})\mathbf{B} = (\mathbf{A}\cdot\mathbf{C})\mathbf{B} \qquad (1.46)$$

represents a vector in the direction of vector **B**.

Similarly

$$(\mathbf{AB})\cdot\mathbf{C} = \mathbf{A}(\mathbf{B}\cdot\mathbf{C}) = (\mathbf{A}\cdot\mathbf{C})\mathbf{B} \qquad (1.47)$$

represents a vector in the direction of vector **A**.

Thus,
$$\mathbf{C}\cdot(\mathbf{AB}) \neq (\mathbf{AB})\cdot\mathbf{C} \qquad (1.48)$$

Some examples of tensors of rank two are: moment of inertia tensor, stress and strain tensor, energy-momentum tensor, etc.

The term dyadic is used in older books and literature and is falling cut of use.

1.11 SCALAR TRIPLE PRODUCT

The products of three vectors are also carried out in two different ways: (i) a scalar triple product $\mathbf{A}\cdot(\mathbf{B} \times \mathbf{C})$, and (ii) a vector triple product $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$.

Consider scalar triple product $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C})$. We can write it as

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = (\mathbf{i}A_x + \mathbf{j}A_y + \mathbf{k}A_z)\cdot\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix}$$

or

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = A_x(B_yC_z - B_zC_y) + A_y(B_zC_x - B_xC_z) + A_z(B_xC_y - B_yC_x) \tag{1.49}$$

or

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix} \tag{1.50}$$

We can regroup the right-hand side of equation (1.49) to give the following combinations,

$$\begin{aligned}\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) &= \mathbf{B}\cdot(\mathbf{C}\times\mathbf{A}) = \mathbf{C}\cdot(\mathbf{A}\times\mathbf{B}) \\ &= -\mathbf{B}\cdot(\mathbf{A}\times\mathbf{C}) = -\,\mathbf{C}\cdot(\mathbf{B}\times\mathbf{A}) = -\,\mathbf{A}\cdot(\mathbf{C}\times\mathbf{B})\end{aligned} \tag{1.51}$$

Equation (1.51) shows that if a cyclic change is made in the sequence of **A**, **B** and **C**, the scalar triple product remains the same. But, if the order is anticyclic, the sign of the product is reversed. We can write equation (1.51) after dropping the parentheses. This is because $\mathbf{A}\cdot\mathbf{B}\times\mathbf{C}$ will always mean product $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C})$. This cannot be looked upon as $(\mathbf{A}\cdot\mathbf{B})\times\mathbf{C}$, because $\mathbf{A}\cdot\mathbf{B}$ is a scalar and its cross product with **C** is meaningless.

Equation (1.51) also shows that the cyclic or anticyclic order of vectors **A**, **B** and **C** is important and not the place of dot or cross. Thus,

$$\mathbf{A}\cdot\mathbf{B}\times\mathbf{C} = \mathbf{A}\times\mathbf{B}\cdot\mathbf{C} \tag{1.52}$$

We come across the scalar triple product quite often and hence for convenience we can drop the dot and cross without any ambiguity. Thus, (**ABC**) will mean $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C})$ and so on.

In this notation, we can write equation (1.51) as,

$$(\mathbf{ABC}) = (\mathbf{BCA}) = (\mathbf{CAB}) = -(\mathbf{ACB}) = -(\mathbf{CBA}) = -(\mathbf{BAC}) \tag{1.53}$$

Each one of these is equal to the right-hand side of equation (1.50).

The scalar triple product has an interesting geometrical interpretation. This can be understood from Fig. 1.16 in which a parallelopiped with sides **A**, **B** and **C** is drawn. We have already remarked that $|\mathbf{B}\times\mathbf{C}| = BC\sin\theta$ represents the area of the parallelogram with sides **B** and **C**. Moreover, area vector $(\mathbf{B}\times\mathbf{C})$ is represented by means of a straight line drawn at right angles to the plane formed by **B** and **C**. Now, the dot product of **A** with $(\mathbf{B}\times\mathbf{C})$ is given by

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = [\text{projection of } \mathbf{A} \text{ on } (\mathbf{B}\times\mathbf{C})] \times [\text{area of the base of the parallelogram}] \tag{1.54}$$

But, the projection of **A** on $(\mathbf{B}\times\mathbf{C})$ is the height or the perpendicular distance between the two opposite faces of the parallelopiped. Thus,

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \text{height of the parallelopiped} \times \text{area of the base of the parallelogram}$$

$$(\mathbf{ABC}) = \text{volume of the parallelopiped} \tag{1.55}$$

If $(\mathbf{ABC}) = 0$, the volume of the parallelopiped is zero. In that case either one or more of the vectors are zero or if all the vectors are non-zero, they are coplanar. Hence, for three non-zero vectors **A**, **B** and **C** to be coplanar, we have the condition

$$(\mathbf{ABC}) = 0 \tag{1.56}$$

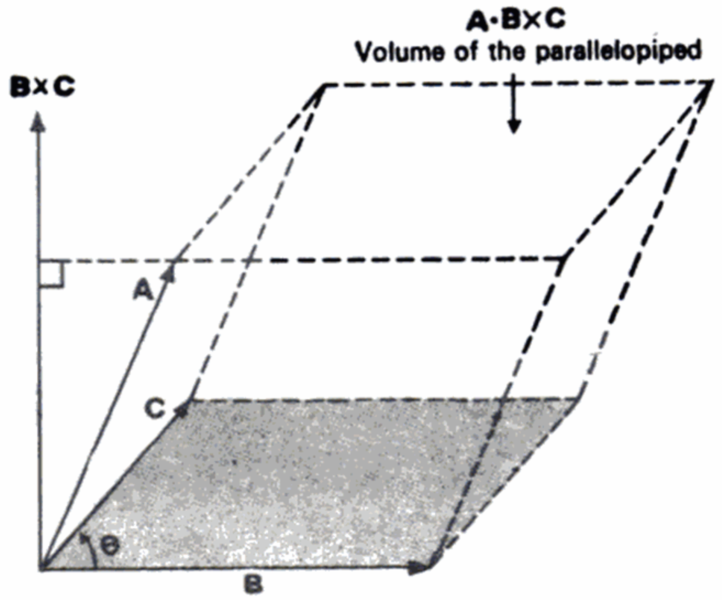


Fig. 1.16 Scalar triple product

Now, any face of the parallelopiped can be taken as the base. Hence, the expressions

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \mathbf{B}\cdot(\mathbf{C}\times\mathbf{A}) = \mathbf{C}\cdot(\mathbf{A}\times\mathbf{B})$$

represent the same volume. In these expressions, the cyclic order is maintained. Hence, the volume of the parallelopiped has a positive sign. If the cyclic order is not maintained, we get a negative sign for the volume. Scalar quantities like this, in which the sign of the quantity depends upon proper cyclic order of the component vectors are referred to as pseudo-scalars. This point is discussed in greater details in article 1.16.

1.12 RECIPROCAL VECTORS

The concept of the reciprocal vectors is very often used in the discussion of reciprocal lattice in solid state physics.

Let us resolve vector **V** into three components along three non-coplanar oblique axes (Fig. 1.17). Let **a**, **b** and **c** form the oblique coordinate axes, which are non-orthogonal. Note that the non-orthogonal base vectors **a**, **b** and **c** need not necessarily be the unit vectors. Such vectors are used in crystallography, in problems dealing with propagation of waves through solids, etc.

Any vector **V** can now be written as

$$\mathbf{V} = V_a\mathbf{a} + V_b\mathbf{b} + V_c\mathbf{c} \tag{1.57}$$

where V_a, V_b and V_c are the components of vector **V** along the directions of **a**, **b** and **c** respectively in the basis of **a**, **b**, **c**. The coordinate axes being non-orthogonal, $V^2 \neq V_a^2 + V_b^2 + V_c^2$.

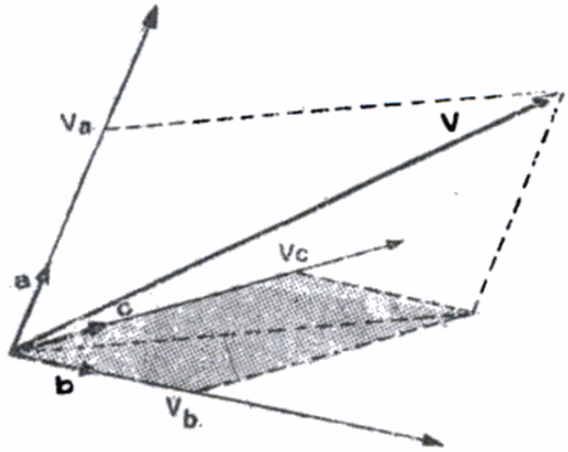


Fig. 1.17 Resolution of a vector along the oblique axes

We now define a set of three vectors **A**, **B** and **C** called reciprocal vectors of the set **a**, **b**, **c** by the expressions

$$\mathbf{A} = \frac{\mathbf{b} \times \mathbf{c}}{(\mathbf{abc})}, \quad \mathbf{B} = \frac{\mathbf{c} \times \mathbf{a}}{(\mathbf{abc})} \quad \text{and} \quad \mathbf{C} = \frac{\mathbf{a} \times \mathbf{b}}{(\mathbf{abc})} \tag{1.58}$$

From the definitions themselves, it can be seen that **A** is perpendicular to the plane formed by **b** and **c** and the magnitude of **A** is proportional to $1/a$.

It is also clear from the definitions themselves that

$$\mathbf{a} \cdot \mathbf{A} = \mathbf{b} \cdot \mathbf{B} = \mathbf{c} \cdot \mathbf{C} = 1 \tag{1.59}$$

and

$$\mathbf{A} \cdot \mathbf{b} = \mathbf{A} \cdot \mathbf{c} = \mathbf{B} \cdot \mathbf{a} = \mathbf{B} \cdot \mathbf{c} = \mathbf{C} \cdot \mathbf{a} = \mathbf{C} \cdot \mathbf{b} = 0 \tag{1.60}$$

Let us now express any vector **P** in terms of reciprocal vectors. This gives

$$\mathbf{P} = P_A\mathbf{A} + P_B\mathbf{B} + P_C\mathbf{C} \tag{1.61}$$

where P_A, P_B and P_C are the components of **P** along the directions of **A**, **B** and **C** respectively.

The dot product of **V** and **P** gives

$$\mathbf{V} \cdot \mathbf{P} = (V_a\mathbf{a} + V_b\mathbf{b} + V_c\mathbf{c}) \cdot (P_A\mathbf{A} + P_B\mathbf{B} + P_C\mathbf{C})$$

or

$$\mathbf{V} \cdot \mathbf{P} = V_aP_A + V_bP_B + V_cP_C \tag{1.62}$$

If, further, $\mathbf{V} \equiv \mathbf{P}$, $\mathbf{V} \cdot \mathbf{P} = V^2$, then equation (1.62) becomes

$$V^2 = V_aV_A + V_bV_B + V_cV_C \tag{1.63}$$

Although we have obtained the same form for the dot product as in equation (1.31), the components of two vectors **V** and **P** are now taken in different coordinate systems.

From the definitions of reciprocal vectors it is obvious that an orthogonal set of unit vectors is its own reciprocal set.

1.13 VECTOR TRIPLE PRODUCT

The vector product of three vectors $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ is written as $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$. In such a product we cannot change the sequence of terms nor the position of parenthesis.

The geometrical aspects of the vector triple product $\mathbf{E} = \mathbf{A} \times (\mathbf{B} \times \mathbf{C})$ are shown in Fig. 1.18. Vector $\mathbf{D} = (\mathbf{B} \times \mathbf{C})$ is perpendicular to the plane formed by vectors $\mathbf{B}$ and $\mathbf{C}$. Hence, vector $\mathbf{E} = \mathbf{A} \times (\mathbf{B} \times \mathbf{C})$ is perpendicular to the plane formed by vectors $\mathbf{A}$ and $\mathbf{D} = (\mathbf{B} \times \mathbf{C})$. Thus, vector $\mathbf{E} = \mathbf{A} \times (\mathbf{B} \times \mathbf{C})$ will lie in the plane formed by vectors $\mathbf{B}$ and $\mathbf{C}$ and further in that plane it will be perpendicular to vector $\mathbf{A}$ (Fig. 1.18).

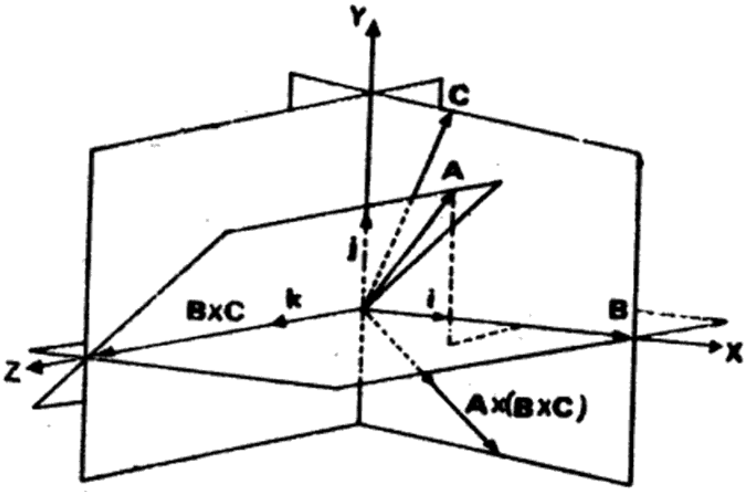


Fig. 1.18 Vector triple product

Similarly, vector $(\mathbf{A} \times \mathbf{B}) \times \mathbf{C}$ will lie in the plane formed by vectors $\mathbf{A}$ and $\mathbf{B}$ and in that plane it will be perpendicular to vector $\mathbf{C}$. Thus,

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) \neq (\mathbf{A} \times \mathbf{B}) \times \mathbf{C} \tag{1.64}$$

Since, in general, vector $\mathbf{E} = \mathbf{A} \times (\mathbf{B} \times \mathbf{C})$ lies in the plane formed by vectors $\mathbf{B}$ and $\mathbf{C}$, we can write

$$\mathbf{E} = \mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = m\mathbf{B} + n\mathbf{C} \tag{1.65}$$

provided that vectors $\mathbf{B}$ and $\mathbf{C}$ are non-collinear. The quantities m and n used in equation (1.65) are scalar quantities to be determined.

But, $\mathbf{E}$ is also perpendicular to $\mathbf{A}$ and hence its dot product with $\mathbf{A}$ will be zero. Thus,

$$\mathbf{A} \cdot \mathbf{E} = 0 = m(\mathbf{A} \cdot \mathbf{B}) + n(\mathbf{A} \cdot \mathbf{C}) \tag{1.66}$$

or

$$\frac{m}{\mathbf{A} \cdot \mathbf{C}} = \frac{-n}{\mathbf{A} \cdot \mathbf{B}} = p \tag{1.67}$$

Now, equation (1.65) can be written as

$$\mathbf{E} = \mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = p\{\mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})\} \tag{1.68}$$

To determine p, let us consider a case in which

$$\mathbf{A} = \mathbf{i}A_x + \mathbf{j}A_y + \mathbf{k}A_z,\ \mathbf{B} = \mathbf{i}B_x \text{ and } \mathbf{C} = \mathbf{i}C_x + \mathbf{j}C_y$$

In this case, we have

$$\begin{aligned}\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) &= (\mathbf{i}A_x + \mathbf{j}A_y + \mathbf{k}A_z) \times \mathbf{k}B_xC_y \\ &= -\,\mathbf{j}A_xB_xC_y + \mathbf{i}A_yB_xC_y \\ &= \mathbf{i}B_x(A_xC_x + A_yC_y) - (\mathbf{i}C_x + \mathbf{j}C_y)(A_xB_x) \\ &= \mathbf{B}(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}(\mathbf{A}\cdot\mathbf{B}) \qquad (1.69)\end{aligned}$$

This shows that $p = 1$ and the vector triple product can be expanded as

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}(\mathbf{A}\cdot\mathbf{B}) \qquad (1.70)$$

This is the most important identity for expansion of the vector triple product. This expansion may be remembered by the 'back-cab' (**BAC** — **CAB**) rule.

The expansion can be better remembered by the statement,

$$\begin{bmatrix}\text{Vector}\\ \text{triple}\\ \text{product}\end{bmatrix} = \begin{pmatrix}\text{Second}\\ \text{vector}\end{pmatrix}\begin{bmatrix}\text{Dot product of}\\ \text{first and}\\ \text{third vectors}\end{bmatrix} - \begin{pmatrix}\text{Third}\\ \text{vector}\end{pmatrix}\begin{bmatrix}\text{Dot product of}\\ \text{first and}\\ \text{second vectors}\end{bmatrix} \qquad (1.71)$$

We can also prove that

$$(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = \mathbf{B}(\mathbf{A}\cdot\mathbf{C}) - \mathbf{A}(\mathbf{B}\cdot\mathbf{C}) \qquad (1.72)$$

The results of equations (1.70) or (1.72) can also be obtained by the straight-forward expansions of the product in terms of the cartesian components.

1.14 ROTATIONAL QUANTITIES AS VECTORS

Two linear displacements $\mathbf{x}_1$ and $\mathbf{x}_2$ can be added vectorially and the vector addition is commutative, i.e., $\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}_2 + \mathbf{x}_1$. The rotation and

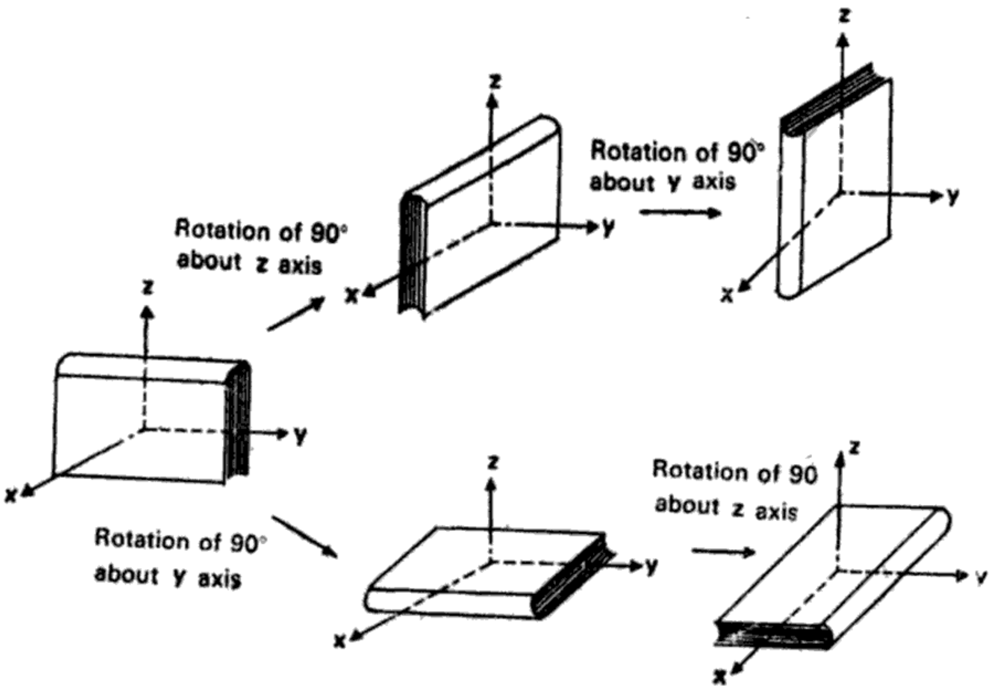


Fig. 1.19 Addition of finite rotations is not commutative

other rotational quantities derived from it may also be treated as vector quantities. The addition of such two vector quantities should, however, be commutative. Fig. 1.19 shows that the addition of two finite rotations, $\theta_1 = 90°$ and $\theta_2 = 90°$, does not yield the same result when the sequence $\theta_2 + \theta_1$ is followed instead of the sequence $\theta_1 + \theta_2$. The body which is subjected to these two finite rotations is not found to be in the same final state. Hence, finite rotations cannot be regarded as vector quantities. If, however, we go on reducing the angle of rotation and rotate the body about the axes whose directions are nearly the same, then we shall see that, for very small rotations $\Delta\theta_1$ and $\Delta\theta_2$, the addition will be commutative.

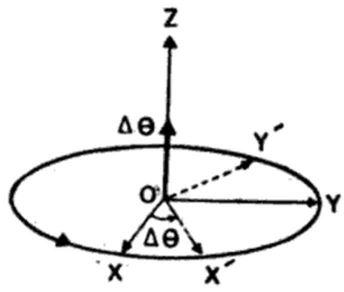


Fig. 1.20 Representation of rotation as a vector

To represent infinitesimal rotation as a vector we draw a straight line along the axis of rotation, (i) the length of which is proportional to the magnitude of the angle of rotation, and (ii) the arrowhead points in the direction of advancement of the tip of the right-hand screw (Fig. 1.20).

Let us now consider some relationships between the angle of rotation and the corresponding linear vector quantities. Suppose that particle P is moving along the circumference of a circle and moves from P to Q in time Δt such that $\angle PCQ = \Delta\theta$. This infinitesimal rotation $\Delta\boldsymbol{\theta}$ about the axis OC (see Fig. 1.21) is represented by vector **CA**. Let $\mathbf{r}$ and $\mathbf{r} + \Delta\mathbf{r}$ be the position vectors of points P and Q with respect to origin O taken on the axis of rotation. Then, linear displacement of the particle is $\mathbf{PQ} = \Delta\mathbf{r}$. Angle of rotation $\Delta\boldsymbol{\theta}$ and linear displacement $\Delta\mathbf{r}$ are related by

$$\Delta\mathbf{r} = \Delta\boldsymbol{\theta} \times \mathbf{r} \tag{1.73}$$

where $\Delta\boldsymbol{\theta}$ is represented by means of a vector as described above.

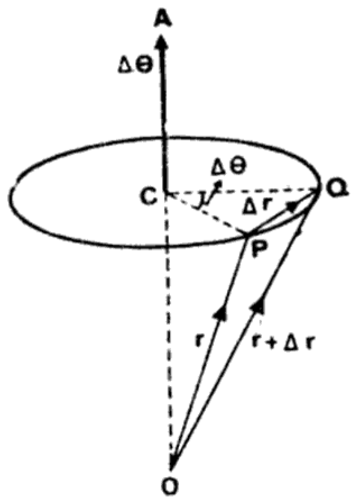


Fig. 1.21 Relation between linear displacement and rotation and angular displacement

Let us now consider two successive rotations $\Delta\boldsymbol{\theta}_1$ and $\Delta\boldsymbol{\theta}_2$ of the particle (not shown in the figure). Here the particle moves through $\Delta\mathbf{r}_1$ from its original position $\mathbf{r}$ by performing rotation $\Delta\theta_1$ followed by second displacement $\Delta\mathbf{r}_2$ from position $\mathbf{r} + \Delta\mathbf{r}_1$ by performing rotation $\Delta\boldsymbol{\theta}_2$. Then, corresponding linear displacements are given by

$$\Delta\mathbf{r}_1 = \Delta\boldsymbol{\theta}_1 \times \mathbf{r}$$

and

$$\Delta\mathbf{r}_2 = \Delta\boldsymbol{\theta}_2 \times (\mathbf{r} + \Delta\mathbf{r}_1)$$

The final position vector after two rotations is given by

$$\mathbf{r} + \Delta\mathbf{r}_{12} = \mathbf{r} + \Delta\mathbf{r}_1 + \Delta\mathbf{r}_2$$
$$= \mathbf{r} + \Delta\boldsymbol{\theta}_1 \times \mathbf{r} + \Delta\boldsymbol{\theta}_2 \times (\mathbf{r} + \Delta\mathbf{r}_1)$$

or
$$\mathbf{r} + \Delta\mathbf{r}_{12} = \mathbf{r} + (\Delta\boldsymbol{\theta}_1 + \Delta\boldsymbol{\theta}_2) \times \mathbf{r} \tag{1.74}$$

after neglecting the second-order infinitesimal quantities. This approximation, in fact, allows us to represent small angular displacement by vectors.

If we reverse the sequence of infinitesimal rotations, i.e. if we displace the particle first through $\Delta\mathbf{r}_2$ from its position $\mathbf{r}$ by rotating through $\Delta\boldsymbol{\theta}_2$ followed by displacement $\Delta\mathbf{r}_1$ from the position $\mathbf{r} + \Delta\mathbf{r}_2$ by rotating through $\Delta\boldsymbol{\theta}_1$, we get

$$\mathbf{r} + \Delta\mathbf{r}_{21} = \mathbf{r} + \Delta\mathbf{r}_2 + \Delta\mathbf{r}_1$$
$$= \mathbf{r} + \Delta\boldsymbol{\theta}_2 \times \mathbf{r} + \Delta\boldsymbol{\theta}_1 \times (\mathbf{r} + \Delta\mathbf{r}_2)$$

or
$$\mathbf{r} + \Delta\mathbf{r}_{21} = \mathbf{r} + (\Delta\boldsymbol{\theta}_2 + \Delta\boldsymbol{\theta}_1) \times \mathbf{r} \tag{1.75}$$

Since, $\Delta\mathbf{r}_{12} = \Delta\mathbf{r}_{21}$, the addition $\Delta\boldsymbol{\theta}_1 + \Delta\boldsymbol{\theta}_2$ is the same as $\Delta\boldsymbol{\theta}_2 + \Delta\boldsymbol{\theta}_1$ Thus, the addition of infinitesimal angular displacements is commutative.

Thus, infinitesimal angular displacement is a vector quantity. This immediately leads us to the definition of angular velocity vector. Dividing angular displacement $\Delta\boldsymbol{\theta}$ by Δt and finding the limit as Δt tends to zero, we get

$$\boldsymbol{\omega} = \lim_{\Delta t \to 0} \frac{\Delta\boldsymbol{\theta}}{\Delta t} = \frac{d\boldsymbol{\theta}}{dt} \tag{1.76}$$

Equation (1.76) states that the angular velocity is the rate of change of angular displacement with respect to time. Angular velocity vector $\boldsymbol{\omega}$ is in the direction of $\Delta\boldsymbol{\theta}$.

Dividing both sides of equation (1.73) by Δt and finding the limit as Δt tends to zero, we get

$$\lim_{\Delta t \to 0} \frac{\Delta\mathbf{r}}{\Delta t} = \lim_{\Delta t \to 0} \frac{\Delta\boldsymbol{\theta}}{\Delta t} \times \mathbf{r}$$

i.e.
$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r} \tag{1.77}$$

It should be noted that other rotational quantities, such as angular momentum **L**, torque **N**, etc., are also regarded as vector quantities. They are represented in a similar manner. These quantities are related to their linear analogues by the relations

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}$$

and
$$\mathbf{N} = \mathbf{r} \times \mathbf{F} \tag{1.78}$$

where the symbols have their usual meanings.

1.15 ROTATION OF COORDINATE AXES

Consider a rotation of the x and y axes about the z axis (Fig. 1.22). Then, any point $P(x, y, z)$ will have to be written as $P(x', y', z')$. The new coordinates of point P will be related to its original coordinates by the relations:

$$\left.\begin{aligned} x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta \\ z' &= z \end{aligned}\right\} \tag{1.79}$$

and

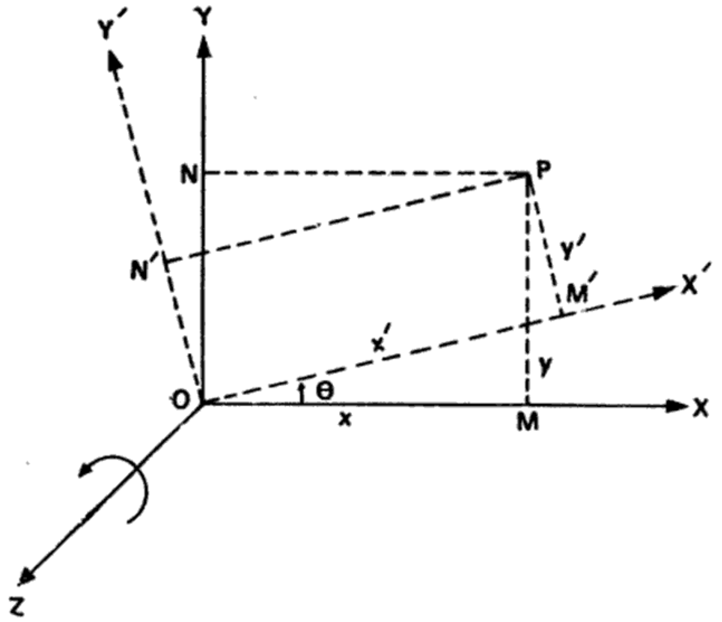


Fig. 1.22 Rotation of the coordinate axes about the z-axis

These relations can be conveniently written in the matrix form

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{1.80}$$

This matrix equation can be written in a compact form by using

$$\mathbf{r}' = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{and} \quad R(\theta) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.81}$$

Then, equation (1.80) becomes

$$\mathbf{r}' = R(\theta)\mathbf{r} \tag{1.82}$$

and represents the change of the position vector from $\mathbf{r}$ to $\mathbf{r}'$ under the rotation of coordinate axes represented by $R(\theta)$. The matrix $R(\theta)$ is, therefore, called a transformation matrix and is regarded as an operator. It operates on $\mathbf{r}$ and changes it to $\mathbf{r}'$.

The transformation of $\mathbf{r}$ to $\mathbf{r}'$ can be obtained in two ways. We can rotate the coordinate axes through angle θ as mentioned above and get $\mathbf{r}'$. In this case, point P is held fixed and the axes are rotated (Fig. 1.22).

We can also transform $\mathbf{r}$ to $\mathbf{r}'$ by keeping the coordinate axes fixed and rotating position vector $\mathbf{r}$ through angle θ in the opposite sense (Fig. 1.23).

Thus, equation (1.82) can also be interpreted as equivalent to rotation of position vector $\mathbf{r}$ to $\mathbf{r}'$ when a matrix operator $R(\theta)$ operates on it.

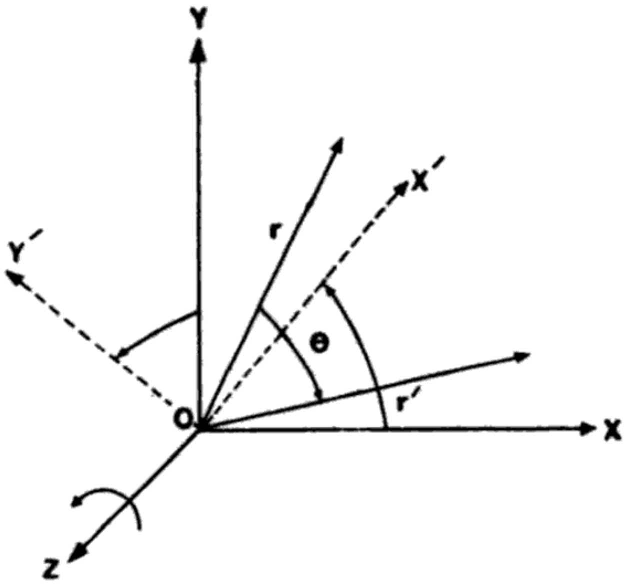


Fig. 1.23 Rotation of position vector **r** in the xy-plane through angle θ is equivalent to rotation of the xy-axes about the z-axis through angle θ in the opposite direction

Operator $R(\theta)$ has some interesting properties. For example, when $\theta = 0°$, i.e. when there is no rotation, the determinant of $R(0°)$ is

$$R\,(\theta = 0°) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \tag{1.83}$$

Now, consider

$$R_1 = \begin{vmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{vmatrix} = -1 \tag{1.84}$$

In that case

$$\mathbf{r}' = R_1\mathbf{r} = -\,\mathbf{r} \tag{1.85}$$

Thus, the effect of operator R_1 is to change the signs of all the components of position vector **r**. Thus $x \to -x$, $y \to -y$ and $z \to -z$. This operation is called inversion or space reflection. In this operation all the coordinate axes are reflected at the origin (Fig. 1.24). The coordinate system obtained by inversion of a right-handed coordinate system is a left-handed coordinate system. It should be noted that even a reflection of one of the axes changes a right-handed coordinate system into a left-handed coordinate system. The two coordinate systems can never be made coincident simply by rotations. Thus, the transition from a right-handed coordinate system to a left-handed coordinate system is not the *physical rotation* in the usual sense of the word. Hence, it is termed as *improper rotation*. Thus, when $R(\theta) = -1$, the rotation is called the *improper rotation*. Naturally when $R(\theta) = 1$, the rotation is a *proper rotation*.

In modern physics, the concept of *parity* is very essential. A physical quantity is said to have *odd parity* if the quantity changes the sign under an inversion of the coordinate system. Conversely, a physical quantity is

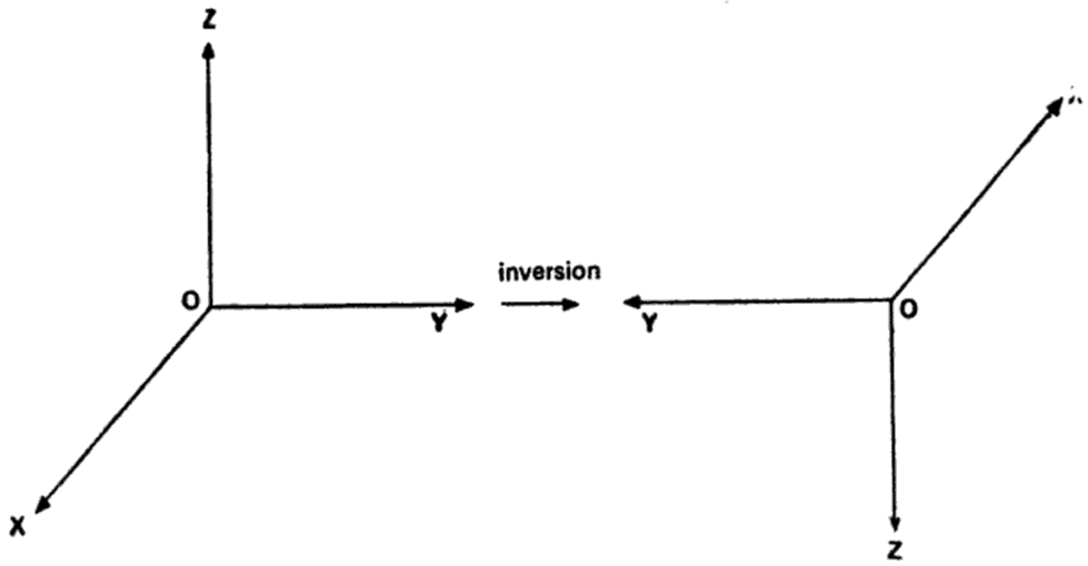


Fig. 1.24 Inversion or space reflection of axes

said to have an *even parity* if the quantity remains invariant under an inversion of the coordinate axes. The quantities like displacement, momentum, force, etc. have odd parity while angular displacement, angular momentum, torque, etc. have even parity.

It can be seen that the inversion of only two axes is a proper rotation. This is because the reflected axes can be made to coincide with the original axes by a rotation of 180° about the third axis.

1.16 PSEUDOVECTORS AND PSEUDOSCALARS

We have come across two types of quantities in Physics. Velocity, acceleration, etc. are quantities wherein direction is clearly indicated by the direction of motion of the particle or the system. Other types of quantities such as angular velocity, angular momentum, etc., are rotational quantities and their direction does not indicate the direction of rotation of the body. A rotating body has a well-defined axis of rotation; however, the sense or direction of rotation, i.e., clockwise or anticlockwise depends on the side the observer looks at the rotation from. In our definition of cross product of two vectors or in representing rotation or area as a vector, we have chosen the right-hand screw rule quite arbitrarily We could have chosen the left-hand screw rule to represent these quantities without affecting the physics of the situation. Use of right-hand screw rule leads us to the right-handed coordinate system, while the left-hand screw rule leads us to the left-handed coordinate system.

We have seen that the behaviour of different physical quantities is different under inversion. The vector quantities like displacement, force,

etc. which change sign under inversion are called the *polar vectors*. Vector quantities, such as angular velocity, torque, etc., remain invariant under inversion. Such vector quantities are called the *axial vectors* or *pseudovectors*.

Consider the reflection of one axis (say the y-axis) of a right-handed coordinate system in the plane of two other axes, viz. the xz-plane. Then we get a left-handed system. This improper rotation is called a *mirror reflection* (Fig. 1.25). In such a case a vector parallel to the y-axis changes the sign while vectors parallel to the x- and z-axes do not change their signs.

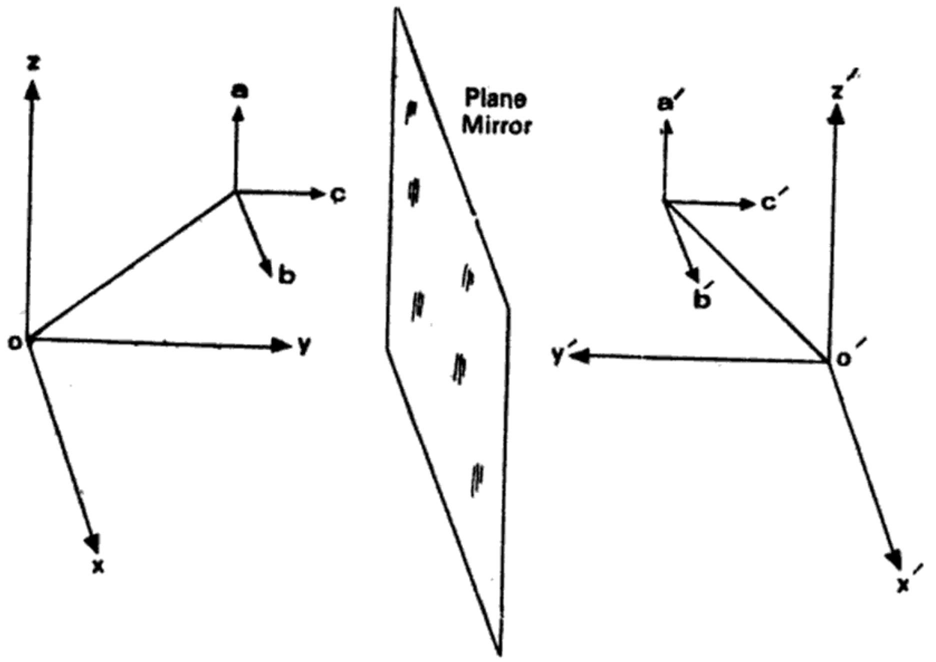


Fig. 1.25 Mirror reflection

Now, let us examine the cross product of **a** and **b** in the mirror reflection. The sense of $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ is obtained by using the right-hand screw rule. The mirror reflection of **a** and **b** is unchanged. Naturally, the cross product of $\mathbf{a}'$ and $\mathbf{b}'$, viz. $\mathbf{c}'$ will also be unchanged. That is why neither angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ nor torque $\mathbf{N} = \mathbf{r} \times \mathbf{F}$ change their sign under a mirror reflection. Thus, the cross product of two polar vectors gives a pseudovector. As the nature of a polar vector is totally different from that of a pseudovector, we do not come across equations in physics in which a polar vector is equated to a pseudovector.

The dot product of vectors **A** and **B** is a scalar quantity, say S. This will remain unchanged under a mirror reflection if both **A** and **B** are polar or axial vectors. Such a scalar quantity is a true scalar. But, quantities such as volume are pseudoscalars as mentioned earlier. Further, if out of **A** and **B**, one is a polar vector and the other is a pseudovector, their dot product S will change sign under a mirror reflection. Such a scalar quantity is called a pseudoscalar. It is obvious that we cannot have expressions equating a scalar with a pseudoscalar.

QUESTIONS

1. Why can an infinitesimal rotation be represented as a vector, whereas a finite rotation cannot ?
2. A rotation is characterised by a magnitude (the angle of rotation) and a direction (axis). Prove that vectors representing finite rotations are not commutative with respect to addition.
3. Suppose that $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$ do not form an orthogonal system. Now show that a component A_i is not the projection of $\mathbf{A}$ on $\hat{\mathbf{e}}_i$ $(i = 1, 2, 3)$.
4. Show that $\mathbf{I} = \mathbf{ii} + \mathbf{jj} + \mathbf{kk}$ works as a unit diad.
5. Why does a mirror reflection reverse left and right (lateral inversion); but leave upward and downward direction unchanged ?
6. Show that the cross product of two pseudovectors is a pseudovector.
7. If l_1, m_1 and n_1 are the direction cosines of $\mathbf{A}$, and l_2, m_2 and n_2 those of $\mathbf{B}$, show that

$$\cos\theta = l_1 l_2 + m_1 m_2 + n_1 n_2$$

   where θ is the angle between $\mathbf{A}$ and $\mathbf{B}$.
8. Why is the scalar product of two vectors $\mathbf{A}$ and $\mathbf{B}$ defined as $AB\cos\theta$ and not by some other expression ?
9. Why is the vector product of two vectors $\mathbf{A}$ and $\mathbf{B}$ defined as $\mathbf{C} = \mathbf{A} \times \mathbf{B} = \hat{\mathbf{n}}\,AB\sin\theta$ and not by some other expression ?
10. Give the geometrical interpretation of scalar triple product.
11. Prove that $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}(\mathbf{A}\cdot\mathbf{B})$ by expanding each vector into its rectangular cartesian components.
12. Show that an orthogonal set of unit vectors is its own reciprocal set.
13. What is meant by proper and improper rotations ? Give examples of each.
14. Explain the terms 'pseudoscalar' and 'pseudovector'. Give illustrations of a true scalar, true vector, pseudoscalar and pseudovector.
15. Explain the term 'parity'.
16. (a) Prove that if $\mathbf{A}$ and $\mathbf{B}$ are non-collinear, then the equation

    $x\mathbf{A} + y\mathbf{B} = 0$ implies $x = y = 0$.

    (b) Prove that, if $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are non-collinear, then the equation

    $x\mathbf{A} + y\mathbf{B} + z\mathbf{C} = 0$ implies $x = y = z = 0$.
17. Show that a volume is a pseudoscalar quantity.
18. If we use the left-hand screw rule to represent a vector obtained by a cross product of two vectors, will the formula for vector triple product change ? Explain.

PROBLEMS

1. The cross product of two vectors is perpendicular to each of these vectors. If $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{D}$, to which of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, is $\mathbf{D}$ perpendicular?
2. Each pair of vectors defines a plane. In which plane do
(i) $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$ and (ii) $(\mathbf{A} \times \mathbf{B}) \times (\mathbf{C} \times \mathbf{D})$ lie?
3. What can you conclude about $\mathbf{A}$ if
(i) $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = 0$ and (ii) $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = 0$?
4. Vectors $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are drawn from a common point and their heads form a plane. Show that $\mathbf{A} \times \mathbf{B} + \mathbf{B} \times \mathbf{C} + \mathbf{C} \times \mathbf{A}$ is perpendicular to this plane.
5. Prove that
$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) + \mathbf{B} \times (\mathbf{C} \times \mathbf{A}) + \mathbf{C} \times (\mathbf{A} \times \mathbf{B}) = 0.$$
6. Prove that
$$\mathbf{A} \times \mathbf{B} \times (\mathbf{C} \times \mathbf{D}) = (\mathbf{B} \cdot \mathbf{D})(\mathbf{A} \times \mathbf{C}) - (\mathbf{B} \cdot \mathbf{C})(\mathbf{A} \times \mathbf{D}).$$
7. Prove that
$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{B} \times \mathbf{C}) \times (\mathbf{C} \times \mathbf{A}) = (\mathbf{ABC})^2.$$
8. Prove that
$$\mathbf{A}(\mathbf{BCD}) - \mathbf{B}(\mathbf{CDA}) + \mathbf{C}(\mathbf{DAB}) - \mathbf{D}(\mathbf{ABC}) = 0.$$
9. Prove that
$$(\mathbf{ABC})(\mathbf{F} \times \mathbf{G}) = \begin{vmatrix} \mathbf{A} & \mathbf{B} & \mathbf{C} \\ \mathbf{F} \cdot \mathbf{A} & \mathbf{F} \cdot \mathbf{B} & \mathbf{F} \cdot \mathbf{C} \\ \mathbf{G} \cdot \mathbf{A} & \mathbf{G} \cdot \mathbf{B} & \mathbf{G} \cdot \mathbf{C} \end{vmatrix}.$$
10. Prove that
$$\mathbf{A} \times [(\mathbf{F} \times \mathbf{B}) \times (\mathbf{G} \times \mathbf{C})] + \mathbf{B} \times [(\mathbf{F} \times \mathbf{C}) \times (\mathbf{G} \times \mathbf{A})] \\ + \mathbf{C} \times [(\mathbf{F} \times \mathbf{A}) \times (\mathbf{G} \times \mathbf{B})] = 0.$$
11. Expand $(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) \times (\mathbf{E} \times \mathbf{F})$.
12. Prove that
$$(\mathbf{A} \times \mathbf{P} \quad \mathbf{B} \times \mathbf{Q} \quad \mathbf{C} \times \mathbf{R}) + (\mathbf{A} \times \mathbf{Q} \quad \mathbf{B} \times \mathbf{R} \quad \mathbf{C} \times \mathbf{P}) \\ + (\mathbf{A} \times \mathbf{R} \quad \mathbf{B} \times \mathbf{P} \quad \mathbf{C} \times \mathbf{Q}) = 0.$$
13. Prove that
$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) + (\mathbf{B} \times \mathbf{C}) \cdot (\mathbf{A} \times \mathbf{D}) - (\mathbf{C} \times \mathbf{A}) \cdot (\mathbf{B} \times \mathbf{D}) = 0.$$
14. Prove that
$$(\mathbf{A} + \mathbf{B} + \mathbf{C}) \times (\mathbf{B} \times \mathbf{C}) \cdot \mathbf{C} = (\mathbf{ABC}).$$
15. Given: $F_x = \frac{y}{r^n}$, $F_y = -\frac{x}{r^n}$, $F_z = 0$, $r = \sqrt{x^2 + y^2}$ and n is a constant, prove that $\mathbf{F}$ represents a tangent to the circle about the origin in the xy-plane.

16. By using the unit vectors

$$\mathbf{p} = \mathbf{i}\cos\theta + \mathbf{j}\sin\theta$$
$$\mathbf{q} = \mathbf{i}\cos\phi - \mathbf{j}\sin\phi$$

and
$$\mathbf{r} = \mathbf{i}\cos\phi + \mathbf{j}\sin\phi,$$

prove the familiar trigonometric results

$$\sin(\theta+\phi) = \sin\theta\cos\phi + \cos\theta\sin\phi$$
$$\cos(\theta+\phi) = \cos\theta\cos\phi - \sin\theta\sin\phi$$
$$\sin(\theta-\phi) = \sin\theta\cos\phi - \cos\theta\sin\phi$$

and
$$\cos(\theta-\phi) = \cos\theta\cos\phi + \sin\theta\sin\phi.$$

17. Derive sine law,

$$\frac{\sin\alpha}{A} = \frac{\sin\beta}{B} = \frac{\sin\gamma}{C},$$

where α, β and γ are respectively the angles opposite to vectors **A**, **B** and **C** which form a triangle.

18. Obtain the cosine law.

19. The Lorentz force due to magnetic induction **B** on charge q moving with velocity **v** is given by

$$\mathbf{F} = q(\mathbf{v}\times\mathbf{B}).$$

The forces for various orientations of the velocity are

$$\text{for } \mathbf{v} = \mathbf{i},\quad \mathbf{F} = q(2\mathbf{k} - 4\mathbf{j})$$
$$\text{for } \mathbf{v} = \mathbf{j},\quad \mathbf{F} = q(4\mathbf{i} - \mathbf{k})$$

and
$$\text{for } \mathbf{v} = \mathbf{k},\quad \mathbf{F} = q(\mathbf{j} - 2\mathbf{i})$$

Find magnetic induction **B** from this data.

20. Show that any arbitrary vector **A** can be expanded as

$$\mathbf{A} = \hat{\mathbf{e}}(\mathbf{A}\cdot\hat{\mathbf{e}}) + \hat{\mathbf{e}}\times(\mathbf{A}\times\hat{\mathbf{e}}),$$

where $\hat{\mathbf{e}}$ is a unit vector in some fixed direction. What is the geometrical significance of the two terms in the expansion?

21. Any vector can be expressed in terms of three non-coplanar vectors **a**, **b** and **c** ($\mathbf{a}\cdot\mathbf{b}\times\mathbf{c} \neq 0$). The reciprocal vectors are defined as

$$\mathbf{A} = \frac{\mathbf{b}\times\mathbf{c}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}},\ \mathbf{B} = \frac{\mathbf{c}\times\mathbf{a}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}} \text{ and } \mathbf{C} = \frac{\mathbf{a}\times\mathbf{b}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}}$$

Show that

(a) $\mathbf{A}\cdot\mathbf{a} = \mathbf{B}\cdot\mathbf{b} = \mathbf{C}\cdot\mathbf{c} = 1$

(b) $\mathbf{A}\cdot\mathbf{b} = \mathbf{A}\cdot\mathbf{c} = \mathbf{B}\cdot\mathbf{a} = \mathbf{B}\cdot\mathbf{c} = \mathbf{C}\cdot\mathbf{a} = \mathbf{C}\cdot\mathbf{b} = 0$

(c) $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \dfrac{1}{\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})}$

(d) $\mathbf{a} = \dfrac{\mathbf{B}\times\mathbf{C}}{\mathbf{A}\cdot\mathbf{B}\times\mathbf{C}}$ etc.

22. Show that

(a) $\sum_{i,j} \mathcal{E}_{ijk}\,\delta_{ij} = 0$

(b) $\sum_{j,k} \mathcal{E}_{ijk}\,\mathcal{E}_{ljk} = 2\delta_{il}$

(c) $\sum_{i,j,k} \mathcal{E}_{ijk}\,\mathcal{E}_{ijk} = 6$

23. Show that

$$(\mathbf{ABC}) = (\mathbf{A}\cdot\mathbf{B}\times\mathbf{C}) = \sum_{i,j,k} \mathcal{E}_{ijk}A_iB_jC_k$$

24. Evaluate the sum

$$\sum_k \mathcal{E}_{ijk}\mathcal{E}_{lmk}$$

by considering the result for all possible combinations of i, j, l, m, viz., $i=j$, $i=l$, $i=m$, $j=l$, $j=m$, $l=m$, $i \neq l$ or m and $j=l$ or m. Note that this contains 81 terms.

Show that

$$\sum_k \mathcal{E}_{ijk}\mathcal{E}_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$$

and use this result to prove that

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = (\mathbf{A}\cdot\mathbf{C})\mathbf{B} - (\mathbf{A}\cdot\mathbf{B})\mathbf{C}$$

25. Use Levi-Civita density to prove the identity

$$(\mathbf{A}\times\mathbf{B})\times(\mathbf{C}\times\mathbf{D}) = \mathbf{C}(\mathbf{ABD}) - \mathbf{D}(\mathbf{ABC})$$

2
Vector Analysis

In this chapter, we shall discuss the operations of differentiation and integration of vectors. These concepts will then be utilized in defining the gradient of a scalar point function and the divergence and curl of a vector point function. Gauss' and Stokes' theorems are then considered. In order to avoid digression in this text book of mechanics, the topic of curvilinear coordinates, which is important in vector analysis, has been included in Appendix A.

2.1 DIFFERENTIATION OF A VECTOR WITH RESPECT TO A SCALAR

Consider a vector **A** which is a continuous function of some scalar variable, say t. Then

$$\mathbf{A} \equiv \mathbf{A}(t) \tag{2.1}$$

As t changes to $t + \Delta t$, the vector undergoes changes in its magnitude and direction. Let this change be denoted by $\mathbf{A} + \Delta\mathbf{A}$. In Fig. 2.1, **OP**

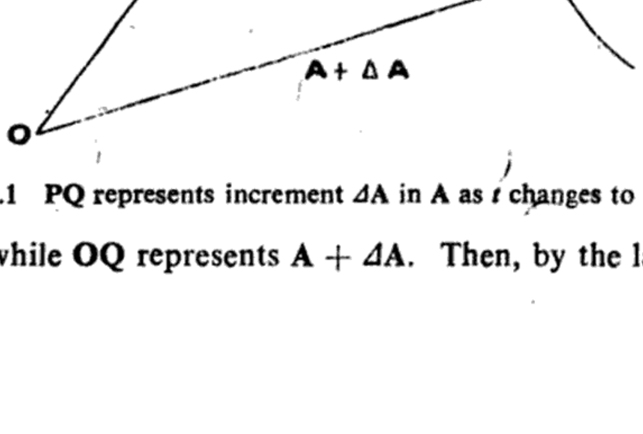


Fig. 2.1 PQ represents increment $\Delta\mathbf{A}$ in **A** as t changes to $t + \Delta t$

represents **A** while **OQ** represents $\mathbf{A} + \Delta\mathbf{A}$. Then, by the law of addition

of vectors, **PQ** represents increment $\Delta\mathbf{A}$ in vector **A** as t changes by amount Δt.

Then, the derivative of vector **A** with respect to scalar t is defined by the equation

$$\frac{d\mathbf{A}}{dt} = \lim_{\Delta t \to 0} \frac{\Delta\mathbf{A}}{\Delta t} = \lim_{\Delta t \to 0} \frac{\mathbf{A}(t + \Delta t) - \mathbf{A}(t)}{\Delta t} \tag{2.2}$$

It is clear from Fig. 2.1 that as t changes, the tip of the vector **A** traces out curve S. Since, $\Delta\mathbf{A}$ is represented by **PQ**, the derivative $\frac{d\mathbf{A}}{dt}$ is a vector whose direction is the limiting direction of $\Delta\mathbf{A}$ as $\Delta t \to 0$. In other words, the direction of $\frac{d\mathbf{A}}{dt}$ lies along the tangent to the curve at point P.

The following equations, which can be proved very easily from the first principles, state the various rules of differentiation of a vector with respect to a scalar variable.

$$\frac{d\mathbf{V}}{dt} = \mathbf{i}\frac{dV_x}{dt} + \mathbf{j}\frac{dV_y}{dt} + \mathbf{k}\frac{dV_z}{dt} \tag{2.3}$$

$$\frac{d}{dt}(c\mathbf{V}) = \frac{dc}{dt}\mathbf{V} + c\frac{d\mathbf{V}}{dt} \tag{2.4}$$

where c is a scalar function of t.

$$\frac{d}{dt}(\mathbf{A} \pm \mathbf{B}) = \frac{d\mathbf{A}}{dt} \pm \frac{d\mathbf{B}}{dt} \tag{2.5}$$

$$\frac{d}{dt}(\mathbf{A}\cdot\mathbf{B}) = \frac{d\mathbf{A}}{dt}\cdot\mathbf{B} + \mathbf{A}\cdot\frac{d\mathbf{B}}{dt} \tag{2.6}$$

Equation (2.6) gives, when $\mathbf{A} = \mathbf{B}$,

$$\frac{d}{dt}(\mathbf{A}\cdot\mathbf{A}) = \frac{d}{dt}(A^2)$$

i.e.

$$2\mathbf{A}\cdot\frac{d\mathbf{A}}{dt} = 2A\frac{dA}{dt}$$

or

$$\mathbf{A}\cdot\frac{d\mathbf{A}}{dt} = A\frac{dA}{dt} \tag{2.7}$$

In a special case, when **A** is a vector of constant magnitude, we have $\frac{dA}{dt} = 0$. Hence, $\mathbf{A}\cdot\frac{d\mathbf{A}}{dt} = 0$. This shows that vector $\frac{d\mathbf{A}}{dt}$ is perpendicular to vector **A**. This situation is observed when a particle moves on the surface of a sphere. In that case, the velocity of the particle, i.e., $\frac{d\mathbf{r}}{dt}$ is always perpendicular to radius vector **r**.

$$\frac{d}{dt}(\mathbf{A}\times\mathbf{B}) = \frac{d\mathbf{A}}{dt}\times\mathbf{B} + \mathbf{A}\times\frac{d\mathbf{B}}{dt}\cdot \tag{2.8}$$

Since the cross product is non-commutative, the sequence of vectors

in each term in equation (2.8) cannot be reversed.

$$\frac{d}{dt}[\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C})]=\frac{d\mathbf{A}}{dt}\cdot(\mathbf{B}\times\mathbf{C})+\mathbf{A}\cdot\left(\frac{d\mathbf{B}}{dt}\times\mathbf{C}\right)+\mathbf{A}\cdot\left(\mathbf{B}\times\frac{d\mathbf{C}}{dt}\right)\quad(2.9)$$

$$\frac{d}{dt}[\mathbf{A}\times(\mathbf{B}\times\mathbf{C})]=\frac{d\mathbf{A}}{dt}\times(\mathbf{B}\times\mathbf{C})+\mathbf{A}\times\left(\frac{d\mathbf{B}}{dt}\times\mathbf{C}\right)+\mathbf{A}\times\left(\mathbf{B}\times\frac{d\mathbf{C}}{dt}\right)\quad(2.10)$$

Thus it is observed that the differentiation in vector analysis follows the same rules as in differential calculus. The only difference is that of the non-commutative property of the vector product of two or more vectors.

2.2 DIFFERENTIATION WITH RESPECT TO TIME—COMPUTATION OF VELOCITY AND ACCELERATION

Consider a particle tracing out some trajectory in space (Fig. 2.2a). Let **OP** = **r** represent the position vector of the particle referred to some

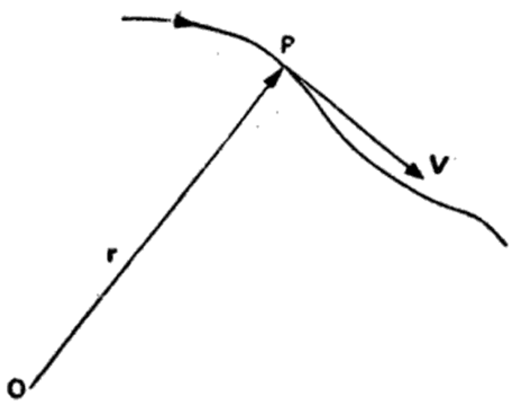


Fig. 2.2a Trajectory of particle *P*

origin *O*. As the particle moves along the curve, position vector **r** is found to change with respect to time. Thus, **r** is a continuous function of time (Fig. 2.2b).

Differentiating **r** w.r.t. *t*, we get velocity **v** as

$$\mathbf{v}=\dot{\mathbf{r}}=\frac{d\mathbf{r}}{dt}\quad(2.11)$$

Further, differentiating **v** w.r.t. *t*, we get acceleration **a** as

$$\mathbf{a}=\dot{\mathbf{v}}=\ddot{\mathbf{r}}=\frac{d^2\mathbf{r}}{dt^2}\quad(2.12)$$

To describe the motion of a particle in the three-dimensional space, we normally use the rectangular cartesian

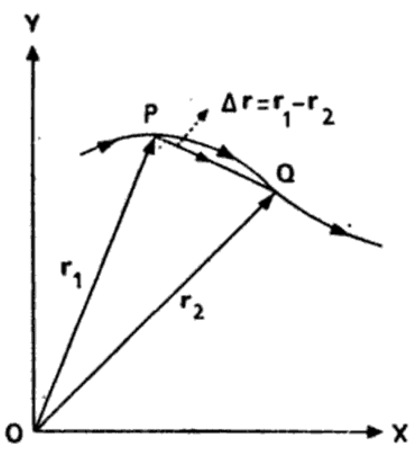


Fig. 2.2b Displacement of a particle from *P* to *Q*

coordinate system. Then,

$$\left.\begin{aligned} \mathbf{r} &= \mathbf{i}x + \mathbf{j}y + \mathbf{k}z \\ \text{Hence,}\quad \mathbf{v} &= \mathbf{i}\dot{x} + \mathbf{j}\dot{y} + \mathbf{k}\dot{z} \\ \text{and}\quad \mathbf{a} &= \mathbf{i}\ddot{x} + \mathbf{j}\ddot{y} + \mathbf{k}\ddot{z} \end{aligned}\right\} \tag{2.13}$$

To describe the motion of a particle in a plane, for instance the motion of a particle in a central force field, it is convenient to use plane polar coordinates. In a central force field, the force acting on the particle is directed along the radius vector towards a fixed point. In the plane polar coordinate system, the position of the particle is expressed in terms of radius vector $\mathbf{r}$ and angle θ that it makes with a fixed direction (say x-axis as in Fig. 2.3). Coordinates r and θ are called the plane polar coordinates of point P and are related to its cartesian coordinates x and y by the formulae,

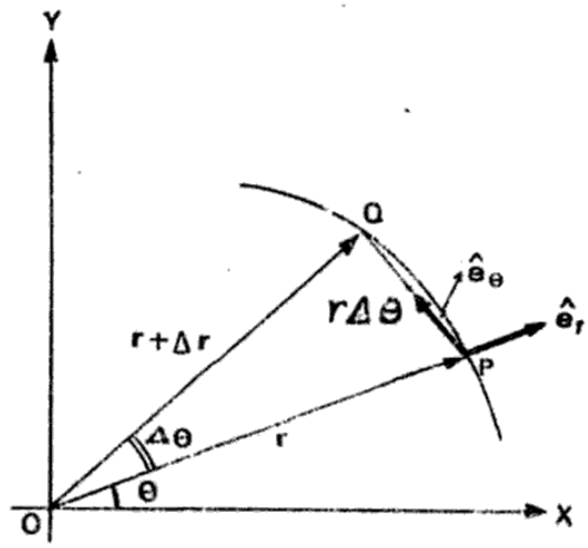


Fig. 2.3 Plane polar coordinates

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta \tag{2.14}$$

The unit vectors for plane polar coordinates are defined by the following general statement. *A unit vector corresponding to a particular coordinate lies along the direction in which position vector* $\mathbf{r}$ *changes when that coordinate is increased by an infinitesimal amount, the other coordinate being unchanged.*

Thus, the unit vectors can be expressed as follows:

$$\hat{\mathbf{e}}_r = \lim_{\Delta r\to 0}\frac{\Delta\mathbf{r}}{\Delta r} = \frac{\partial\mathbf{r}}{\partial r} \qquad (\theta = \text{const}) \tag{2.15a}$$

$$\hat{\mathbf{e}}_\theta = \lim_{\Delta\theta\to 0}\frac{\Delta\mathbf{r}}{r\Delta\theta} = \frac{1}{r}\frac{\partial\mathbf{r}}{\partial\theta} \quad (r = \text{const}) \tag{2.16a}$$

Equations (2.15a) and (2.16a) can also be written as

$$\hat{\mathbf{e}}_r = \frac{\partial\mathbf{r}}{\partial r}\bigg/\left|\frac{\partial\mathbf{r}}{\partial r}\right| \tag{2.15b}$$

and $$\hat{\mathbf{e}}_\theta = \frac{\partial \mathbf{r}}{\partial \theta} \Big/ \left| \frac{\partial \mathbf{r}}{\partial \theta} \right| \tag{2.16b}$$

respectively.

Position vector $\mathbf{r}$ can be written in terms of plane polar coordinates as,

$$\mathbf{r} = r\hat{\mathbf{e}}_r \tag{2.17}$$

Unit vectors $\hat{\mathbf{e}}_r$ and $\hat{\mathbf{e}}_\theta$ are not constant vectors because the directions of these vectors change continuously as the particle moves along the curve.

We can express position vector $\mathbf{r}$ in terms of its cartesian components x and y as

$$\mathbf{r} = \mathbf{i}x + \mathbf{j}y$$

From Fig. 2.3 and equation (2.14), we have

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta$$

Hence, $$\mathbf{r} = \mathbf{i}\,r\cos\theta + \mathbf{j}\,r\sin\theta \tag{2.18}$$

Using equations (2.15a), (2.16a) and (2.18), we get

$$\hat{\mathbf{e}}_r = \mathbf{i}\cos\theta + \mathbf{j}\sin\theta \tag{2.19}$$

and $$\hat{\mathbf{e}}_\theta = -\mathbf{i}\sin\theta + \mathbf{j}\cos\theta \tag{2.20}$$

Equations (2.19) and (2.20) show that both $\hat{\mathbf{e}}_r$ and $\hat{\mathbf{e}}_\theta$ are functions of θ. Further, $\hat{\mathbf{e}}_r \cdot \hat{\mathbf{e}}_\theta = 0$. From this we see that $\hat{\mathbf{e}}_r$ and $\hat{\mathbf{e}}_\theta$ must be perpendicular to each other since $\hat{\mathbf{e}}_r \neq 0$ and $\hat{\mathbf{e}}_\theta \neq 0$. Thus, the plane polar coordinates form an orthogonal coordinate system. Here vector $\hat{\mathbf{e}}_r$ is radial and $\hat{\mathbf{e}}_\theta$ is transverse.

As the position of a particle changes with time, unit vectors $\hat{\mathbf{e}}_r$ and $\hat{\mathbf{e}}_\theta$ must also change with time. Differentiating equations (2.19) and (2.20) with respect to θ, we get

$$\frac{d\hat{\mathbf{e}}_r}{d\theta} = -\mathbf{i}\sin\theta + \mathbf{j}\cos\theta = \hat{\mathbf{e}}_\theta \tag{2.21}$$

and $$\frac{d\hat{\mathbf{e}}_\theta}{d\theta} = -\mathbf{i}\cos\theta - \mathbf{j}\sin\theta = -\hat{\mathbf{e}}_r \tag{2.22}$$

Differentiating equation (2.17) with respect to time we get

$$\mathbf{v} = \dot{\mathbf{r}} = \frac{d\mathbf{r}}{dt} = \frac{d}{dt}(r\hat{\mathbf{e}}_r) = \frac{dr}{dt}\hat{\mathbf{e}}_r + r\frac{d\hat{\mathbf{e}}_r}{dt}$$

$$= \frac{dr}{dt}\hat{\mathbf{e}}_r + r\frac{d\hat{\mathbf{e}}_r}{d\theta}\frac{d\theta}{dt}$$

Hence, by using equation (2.21),

$$\mathbf{v} = \dot{r}\hat{\mathbf{e}}_r + r\dot{\theta}\hat{\mathbf{e}}_\theta \tag{2.23}$$

This shows that the *radial* component of the velocity is given by $v_r = \dot{r}$ while the *transverse* or *angular* component of the velocity is given by $v_\theta = r\dot{\theta}$. Component v_θ arises on account of the change in the direction of position vector $\mathbf{r}$.

Differentiating equation (2.23) with respect to time, we get,

$$\mathbf{a} = \dot{\mathbf{v}} = \frac{d^2\mathbf{r}}{dt^2} = \ddot{r}\hat{\mathbf{e}}_r + \dot{r}\frac{d\hat{\mathbf{e}}_r}{dt} + \dot{r}\dot{\theta}\hat{\mathbf{e}}_\theta + r\ddot{\theta}\hat{\mathbf{e}}_\theta + r\dot{\theta}\frac{d\hat{\mathbf{e}}_\theta}{dt}$$

$$= \ddot{r}\hat{\mathbf{e}}_r + \dot{r}\frac{d\hat{\mathbf{e}}_r}{d\theta}\frac{d\theta}{dt} + \dot{r}\dot{\theta}\hat{\mathbf{e}}_\theta + r\ddot{\theta}\hat{\mathbf{e}}_\theta + r\dot{\theta}\frac{d\hat{\mathbf{e}}_\theta}{d\theta}\frac{d\theta}{dt}$$

$$= \ddot{r}\hat{\mathbf{e}}_r + \dot{r}\dot{\theta}\hat{\mathbf{e}}_\theta + \dot{r}\dot{\theta}\hat{\mathbf{e}}_\theta + r\ddot{\theta}\hat{\mathbf{e}}_\theta - r\dot{\theta}^2\hat{\mathbf{e}}_r$$

or

$$\mathbf{a} = (\ddot{r} - r\dot{\theta}^2)\hat{\mathbf{e}}_r + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\hat{\mathbf{e}}_\theta \tag{2.24}$$

From equation (2.24), we get the radial and transverse acceleration as

$$a_r = \ddot{r} - r\dot{\theta}^2 \quad \text{and} \quad a_\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta}$$

The term $r\dot{\theta}^2 = (v^2/r)$ represents the centripetal acceleration of a particle performing a uniform circular motion. In such a motion, (i) $r =$ constant, and hence $\dot{r} = 0$ and $\ddot{r} = 0$, and (ii) $\dot{\theta} =$ constant and hence $\ddot{\theta} = 0$. Thus, for a particle performing uniform circular motion,

$$a_r = -r\dot{\theta}^2 = -v^2/r \tag{2.25}$$

is the centripetal acceleration of the particle and there is no transverse acceleration.

The calculations of velocity and acceleration in terms of the spherical polar coordinates and cylindrical polar coordinates are given in Appendix A.

2.3 INTEGRATION OF VECTORS

The usual procedures of integral calculus can be directly applied to vector integration. In general, a vector integral can be converted into scalar integrals and, in turn, these can be evaluated by the usual methods.

(a) Line Integral

In physics, we often come across integrals of the type

$$\int_c \Phi\, d\mathbf{r}, \int_c \mathbf{V}\cdot d\mathbf{r} \text{ and } \int_c \mathbf{V} \times d\mathbf{r}$$

where $\Phi = \Phi(x, y, z)$ is a scalar point function representing a scalar field, while $\mathbf{V} = \mathbf{V}(x, y, z)$ is a vector point function representing a vector field and c is some contour along which the integration is to be carried out. We can write these integrals as,

$$\int_c \Phi\, d\mathbf{r} = \mathbf{i}\int_c \Phi(x, y, z)\, dx + \mathbf{j}\int_c \Phi(x, y, z)\, dy + \mathbf{k}\int \Phi(x, y, z)\, dz \tag{2.26}$$

$$\int_c \mathbf{V}\cdot d\mathbf{r} = \int_c V_x(x, y, z)\, dx + \int_c V_y(x, y, z)\, dy + \int V_z(x, y, z)\, dz \tag{2.27}$$

and

$$\int_c \mathbf{V}\times d\mathbf{r} = \mathbf{i}\int_c [V_y(x, y, z)\, dz - V_z(x, y, z)\, dy]$$
$$+ \mathbf{j}\int_c [V_z(x, y, z)\, dx - V_x(x, y, z)\, dz]$$
$$+ \mathbf{k}\int_c [V_x(x, y, z)\, dy - V_y(x, y, z)\, dx] \tag{2.28}$$

Note that **i, j** and **k** are constant unit vectors, i.e. their magnitude and directions do not depend upon position vector **r** and hence these are taken out of integration sign. The integrals on the right-hand sides of equations (2.26), (2.27) and (2.28) can be evaluated by the usual rules of integration. Thus, the integral with respect to x can be evaluated only if we know the dependence of y and z on x. In other words, contour c along which the integration is to be carried out must be known.

If, in equation (2.26), $\Phi = 1$, then the integral

$$\int_c \Phi \, d\mathbf{r} = \int_c d\mathbf{r}$$

simply represents the displacement from A to B where A and B are the starting and end points, respectively, of curve c (Fig. 2.4).

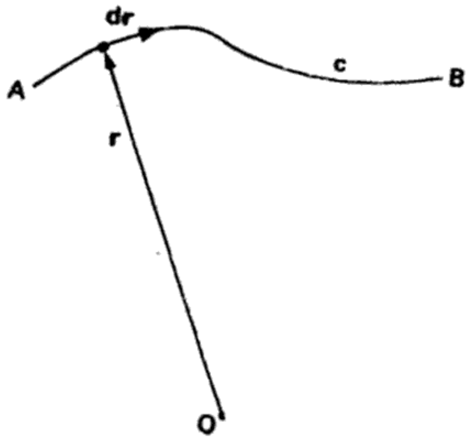


Fig. 2.4 Path of a particle at **r** along curve c from A to B

(b) Surface Integral

We have already seen that a plane surface area can be represented as a vector quantity. The area vector is represented by a directed straight line at right angles to the plane of the area and its sense depends upon the sense in which the bounding curve is described (Fig. 2.5). In physics, we come across surface integrals of the type

$$\int_\sigma \Phi \, d\boldsymbol{\sigma}, \int_\sigma \mathbf{V} \cdot d\boldsymbol{\sigma}, \text{ and } \int_\sigma \mathbf{V} \times d\boldsymbol{\sigma}$$

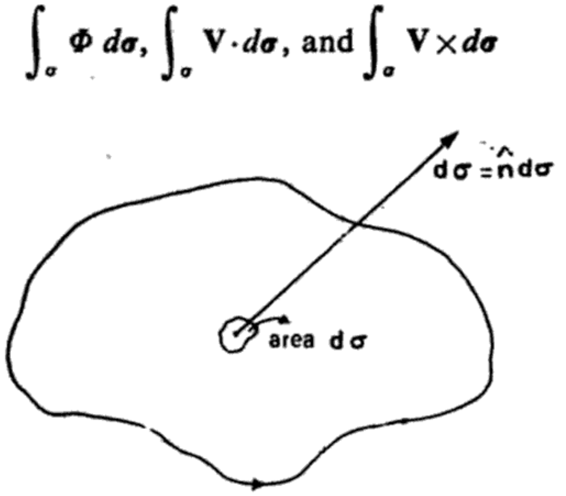


Fig. 2.5 Plane surface area as a vector

where Φ and $\mathbf{V}$ are scalar and vector point functions respectively, and $d\boldsymbol{\sigma}$ is an element of the given surface σ over which the integrals are to be evaluated. As in the case of line integrals, these integrals are written in the scalar form and then evaluated term by term by applying the usual rules of integration.

The surface integral $\int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma}$ is interpreted as a flow or flux of vector $\mathbf{V}$ through surface σ. In order to be able to evaluate these surface integrals, the nature of the surface needs to be specified.

Illustration: To evaluate $\int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma}$ over the surface of a cylinder (Fig. 2.6) $x^2 + y^2 = a^2$ and $z = h$ if $\mathbf{V} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}z$.

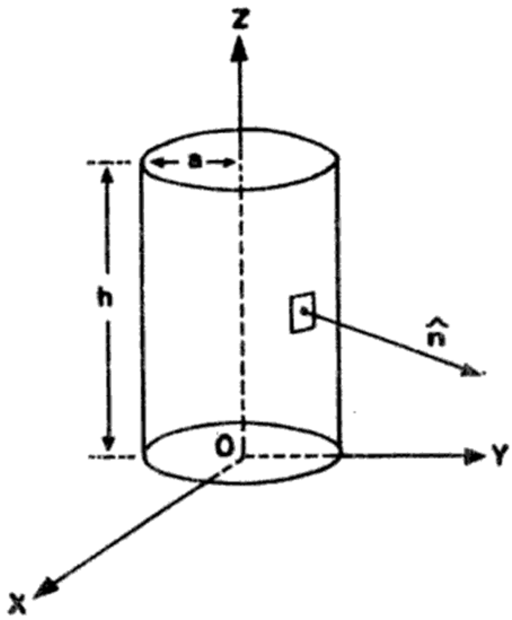


Fig. 2.6 Cylinder $x^2 + y^2 = a^2$ of height h

The unit normal vector on the curved surface is given by

$$\hat{\mathbf{n}} = \frac{\mathbf{i}x + \mathbf{j}y}{\sqrt{x^2 + y^2}}$$

Hence, $$\mathbf{V}\cdot\hat{\mathbf{n}} = \frac{x^2 + y^2}{\sqrt{x^2 + y^2}} = \sqrt{x^2 + y^2} = a$$

Hence, the surface integral over the curved surface is

$$\int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma} = \int_\sigma \mathbf{V}\cdot\hat{\mathbf{n}}\, d\sigma = a\int_\sigma d\sigma = a \times 2\pi a h = 2\pi a^2 h$$

To find the contribution of the top and the bottom surfaces to the total surface integral we note that,

for the bottom surface, $\hat{\mathbf{n}} = -\mathbf{k}$ and $z = 0$

and for the top surface, $\hat{\mathbf{n}} = \mathbf{k}$ and $z = h$

Thus, $\int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma} = -\int_\sigma (\mathbf{i}x + \mathbf{j}y)\cdot\mathbf{k}\, d\sigma = 0$, for the bottom surface,

and $\int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma} = \int_\sigma \mathbf{V}\cdot\mathbf{k}\, d\sigma = h\int_\sigma d\sigma = h\pi a^2$, for the top surface.

Hence, $\int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma} = 3\pi a^2 h$ over the entire surface of the cylinder.

(c) Volume Integral

A volume integral of a vector point function is simpler to evaluate since volume element $d\tau$ (sometimes written as d^3r or d^3x or dv) is itself a scalar quantity.

Thus, the volume integral of vector $\mathbf{V}$ over volume τ is written as

$$\int_\tau \mathbf{V}\, d\tau = \mathbf{i}\int_\tau V_x\, d\tau + \mathbf{j}\int_\tau V_y\, d\tau + \mathbf{k}\int_\tau V_z\, d\tau \tag{2.29}$$

2.4 PARTIAL DIFFERENTIATION

Consider vector $\mathbf{V}$ which is a function of cartesian coordinates x, y and z of a point in space. If y and z remain constant while x increases, we can find partial derivative $\frac{\partial \mathbf{V}}{\partial x}$ which represents the rate of increase of $\mathbf{V}$ with respect to x when variation of y and z is absent. Similarly, $\frac{\partial \mathbf{V}}{\partial y}$ and $\frac{\partial \mathbf{V}}{\partial z}$ denote the partial derivatives of vector $\mathbf{V}$ with respect to y and z respectively. If now x, y and z change simultaneously and if dx, dy and dz denote the differential increments in x, y and z, respectively, the total change or total differential of vector $\mathbf{V}$ is given by

$$d\mathbf{V} = \frac{\partial \mathbf{V}}{\partial x}dx + \frac{\partial \mathbf{V}}{\partial y}dy + \frac{\partial \mathbf{V}}{\partial z}dz \tag{2.30}$$

Let $\mathbf{r} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}z$ be the radius vector from the origin, then its differential increment is

$$d\mathbf{r} = \mathbf{i}\, dx + \mathbf{j}\, dy + \mathbf{k}\, dz \tag{2.31}$$

Equation (2.30) can be written as

$$d\mathbf{V} = \left[dx\frac{\partial}{\partial x} + dy\frac{\partial}{\partial y} + dz\frac{\partial}{\partial z}\right]\mathbf{V} \tag{2.32}$$

If we now define a vector differential operator by the formula

$$\boldsymbol{\nabla} = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z} \tag{2.32a}$$

then operator $\left[dx\frac{\partial}{\partial x} + dy\frac{\partial}{\partial y} + dz\frac{\partial}{\partial z}\right]$ of equation (2.32) will be defined as the dot product of $d\mathbf{r}$ and $\boldsymbol{\nabla}$ (read as *del*). Hence,

$$d\mathbf{V} = (d\mathbf{r}\cdot\boldsymbol{\nabla})\,\mathbf{V} \tag{2.33}$$

The operator $\boldsymbol{\nabla}$ is a vector differential operator and it can operate on a scalar point function or a vector point function.

2.5 GRADIENT OF A SCALAR POINT FUNCTION

Consider a scalar point function $\Phi = \Phi(x, y, z)$ in some region. In this region, we can always draw surfaces over which the magnitude of the scalar point function remains constant.

Consider two neighbouring surfaces such that the value of Φ changes by amount $d\Phi$ as we go from the first to the second surface (Fig. 2.7).

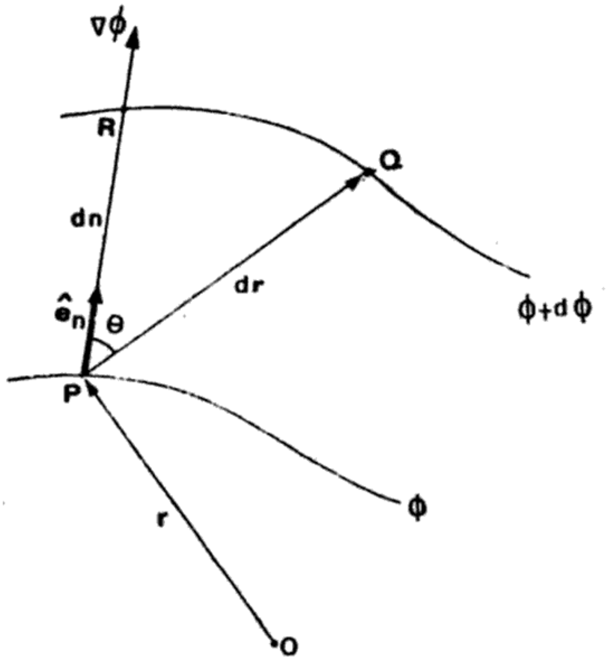


Fig. 2.7 Surfaces Φ = constant and $\Phi + d\Phi$ = constant with normal direction $\hat{\mathbf{e}}_n$ at P

Let $\mathbf{r}$ be the position vector of point P, and $\mathbf{r} + d\mathbf{r}$ that of point Q. Then, the shortest distance between the two surfaces will be

$$PR = dn = dr \cos\theta \tag{2.34}$$

The rate of change of the magnitude of Φ in direction **PQ** is given by $\frac{\partial \Phi}{\partial r}$, when the two surfaces are very close to each other. This rate has maximum value $\frac{\partial \Phi}{\partial n}$ in the direction of **PR**, i.e., of the unit normal vector $\hat{\mathbf{e}}_n$. It is obvious from equation (2.34) that

$$\frac{\partial \Phi}{\partial r} = \frac{\partial \Phi}{\partial n} \cos\theta \tag{2.34a}$$

Thus, the maximum rate of change of Φ is found in the direction of $\hat{\mathbf{e}}_n$ and is a vector denoted by $\hat{\mathbf{e}}_n \frac{\partial \Phi}{\partial n}$. This vector is termed *gradient* of scalar point function $\Phi(x, y, z)$ at point P and is written as

$$\text{grad } \Phi = \hat{\mathbf{e}}_n \frac{\partial \Phi}{\partial n} \tag{2.34b}$$

Hence, the gradient of a scalar field is a vector field, the vector at any point having a magnitude equal to the maximum rate of increase of Φ at

that point and its direction is perpendicular to the surface $\Phi =$ constant at that point.

If Φ represents electrostatic potential, then the electric force on a unit charge at any point is in the direction of the maximum rate of decrease of potential, i.e. normal to equipotential surfaces. The magnitude of this force is equal to the space rate of decrease of potential.

We now prove that grad $\Phi = \nabla\Phi$. For this, we write

$$\nabla\Phi = \mathbf{i}\,\frac{\partial\Phi}{\partial x} + \mathbf{j}\,\frac{\partial\Phi}{\partial y} + \mathbf{k}\,\frac{\partial\Phi}{\partial z}$$

Quantities $\frac{\partial\Phi}{\partial x}, \frac{\partial\Phi}{\partial y}$ and $\frac{\partial\Phi}{\partial z}$ are the rates of increase of Φ in the x, y and z directions respectively. Then, they are the components of a vector having a magnitude and direction of the maximum rate of increase of Φ.

Now,
$$\text{grad } \Phi = \hat{\mathbf{e}}_n \frac{\partial\Phi}{\partial n}$$

On taking the dot product of the two sides of this equation with $d\mathbf{r}$, we get

$$\begin{aligned}(\text{grad } \Phi)\cdot d\mathbf{r} &= \hat{\mathbf{e}}_n \frac{\partial\Phi}{\partial n}\cdot d\mathbf{r} \\ &= \frac{\partial\Phi}{\partial n}\, dr \cos\theta \\ &= \frac{\partial\Phi}{\partial n}\, dn = d\Phi \qquad (2.35)\end{aligned}$$

by equation (2.34).

But,

$$d\Phi = \frac{\partial\Phi}{\partial x}\, dx + \frac{\partial\Phi}{\partial y}\, dy + \frac{d\Phi}{\partial z}\, dz$$

Hence,

$$\begin{aligned}(\text{grad } \Phi)\cdot d\mathbf{r} &= \frac{\partial\Phi}{\partial x}\, dx + \frac{\partial\Phi}{\partial y}\, dy + \frac{\partial\Phi}{\partial z}\, dz \\ &= (\nabla\Phi)\cdot d\mathbf{r}\end{aligned}$$

or
$$\text{grad } \Phi = \nabla\Phi \qquad (2.36)$$

as $d\mathbf{r}$ can be chosen arbitrarily.

Thus, when ∇ operates on a scalar point function Φ we get a vector which is the gradient of Φ.

It does not necessarily follow that all vector fields can be expressed as the gradient of a scalar function. Let us find the condition under which this is possible.

Let $\mathbf{V}$ be a vector which is expressed in terms of scalar Φ by the equation

$$\mathbf{V} = \text{grad } \Phi = \nabla\Phi$$

Now consider two points A and B in this vector field (Fig. 2.8). Draw any path such as that marked by path 1 and consider a small element $d\mathbf{l}$ of this path. Then, the line integral of vector $\mathbf{V}$ for path 1, i.e. from A to B, is given by

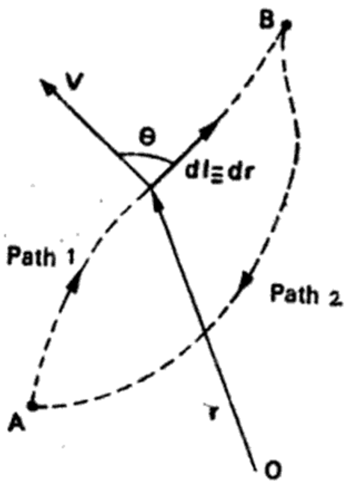


Fig. 2.8 Two paths in a vector field V

$$\int_A^B \mathbf{V}\cdot d\mathbf{l} = \int_A^B (\text{grad } \Phi)\cdot d\mathbf{r}$$
$$= \int_A^B d\Phi = \Phi_B - \Phi_A \qquad (2.37)$$

From Fig. 2.8, it is clear that $d\mathbf{l} \equiv d\mathbf{r}$. In equation (2.37), Φ_B and Φ_A are the values of scalar Φ at points B and A respectively. Since the values of scalar Φ at end points A and B are fixed, the line integral has the same value for various independent paths. Thus, for path 2,

$$\int_B^A \mathbf{V}\cdot d\mathbf{r} = \int_B^A \mathbf{V}\cdot d\mathbf{l} = \Phi_A - \Phi_B \qquad (2.38)$$

Hence, for a closed path ABA, we have

$$\oint \mathbf{V}\cdot d\mathbf{r} = 0$$

i.e.

$$\oint (\text{grad } \Phi)\cdot d\mathbf{r} = 0$$

Thus, when a vector field is expressible as a gradient of a scalar field, the line integral of the vector taken between any two points is independent of the path followed and is equal to the difference between the values of the scalar function at the ends of the path, and further, the line integral around any closed path in such a vector field is zero.

Vector $\mathbf{V}$ is also called a lamellar vector because the field is divided into layers or laminae over which the value of scalar Φ is constant. The field is also called an irrotational field as the line integral of the corresponding vector round any closed path is zero. This point will be described in a later section.

2.6 DIVERGENCE OF A VECTOR

Consider a small volume element $d\tau = dx\ dy\ dz$ as shown in Fig. 2.9. Let $\mathbf{V}$ be a vector point function at point O. Let V_x, V_y and V_z be the components of $\mathbf{V}$ along the x, y and z axes respectively. Similarly, let $\frac{\partial V_x}{\partial x}$, $\frac{\partial V_y}{\partial y}$ and $\frac{\partial V_z}{\partial z}$ be the rates of change of these components in their own directions.

Consider the surfaces marked by 1 and 2. These surfaces are perpendicular to the y axis. The value of y components of $\mathbf{V}$ over these surfaces

are V_y and $V_y + \frac{\partial V_y}{\partial y} dy$. These values can be assumed to remain constant over the respective surfaces as they are very small.

Vector **V** can be ascribed any physical meaning. For instance, let $\mathbf{V} = \rho \mathbf{v}$, where ρ is the density and **v** is the velocity of a fluid. Then $\mathbf{V} = \rho \mathbf{v}$ represents the rate at which the mass of the fluid that flows across a unit area normal to **v** around the point. This flow is very often called the flux of vector **V**.

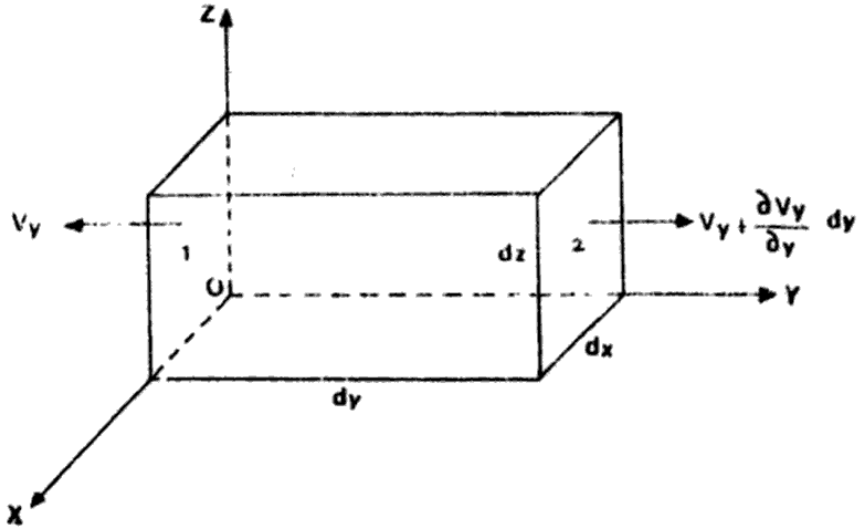


Fig. 2.9 Volume element in vector field **V** at point O

Hence, the net outflow or the flux of vector **V** in the y direction is given by,

$$\left(V_y + \frac{\partial V_y}{\partial y} dy\right) dx\, dz - V_y\, dx\, dz = \frac{\partial V_y}{\partial y} dx\, dy\, dz \tag{2.39}$$

Similarly, the net outflow or the flux of vector **V** in the x and z directions are given by, $\frac{\partial V_x}{\partial x} dx\, dy\, dz$ and $\frac{\partial V_z}{\partial z} dx\, dy\, dz$ respectively. Hence, total outflow or the flux of vector **V** from the volume element

$$= \left(\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}\right) dx\, dy\, dz$$

$$= \left(\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}\right) d\tau \tag{2.40}$$

Thus, the amount of outward flux of vector **V** per unit volume

$$= \left(\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}\right) \tag{2.41}$$

This quantity is referred to as the *divergence* of vector **V**.

It is clear that

$$\text{div } \mathbf{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}$$

$$\text{div } \mathbf{V} = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right)\cdot(\mathbf{i}V_x + \mathbf{j}V_y + \mathbf{k}V_z)$$

or $$\text{div } \mathbf{V} = \nabla\cdot\mathbf{V} \tag{2.42}$$

A positive value of div **V** indicates a net outflow while a negative value of div **V** indicates a net inflow of the flux of vector **V**.

If div $\mathbf{V} = 0$, then there is no net inflow or outflow of the fluid. In other words, the inflow and outflow balance each other. A vector which satisfies this condition is called a *solenoidal* vector. For example, magnetic induction **B** for which $\nabla\cdot\mathbf{B} = 0$ is a solenoidal vector.

Equation (2.40) can be written as

$$\left(\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}\right) dx\, dy\, dz = (\nabla\cdot\mathbf{V})\, d\tau \tag{2.43}$$

This equation can be expressed in the integral form. Consider area $d\boldsymbol{\sigma} = \hat{\mathbf{n}}\, d\sigma$. Then, the outflow of the flux across this area is $\mathbf{V}\cdot\hat{\mathbf{n}}\, d\sigma$. If this element of area is taken on the surface which encloses a certain volume (say the volume element of Fig. 2.9), the net outflow through the entire surface is

$$\int_\sigma \mathbf{V}\cdot\hat{\mathbf{n}}\, d\sigma \tag{2.44}$$

This is the flux of vector **V** through the entire surface of volume element $d\tau$ and hence must be equal to $\nabla\cdot\mathbf{V}\, d\tau$.

Hence $$\nabla\cdot\mathbf{V}\, d\tau = \int_\sigma \mathbf{V}\cdot\hat{\mathbf{n}}\, d\sigma$$

or $$\nabla\cdot\mathbf{V} = \lim_{d\tau\to 0} \frac{\int_\sigma \mathbf{V}\cdot\hat{\mathbf{n}}\, d\sigma}{d\tau} \tag{2.45}$$

2.7 THE EQUATION OF CONTINUITY

Let us apply the above derivation to the flow of a fluid. When

$$\text{div } \mathbf{V} = \text{div } \rho\mathbf{v} = 0 \tag{2.46}$$

the net quantity of fluid in volume $d\tau$ is constant. In this case density ρ is constant and the matter is said to be conserved.

On account of the outflow of the fluid, the density of the fluid inside the volume will decrease. Then, according to the principle of conservation of the total mass of the fluid, we must have

$$\nabla\cdot\rho\mathbf{v}\, d\tau = -\frac{d\rho}{dt}\, d\tau \tag{2.47}$$

where the left-hand side is the loss of mass of this fluid flowing out and the right-hand side is the decrease (negative sign) in the mass due to change of density of the fluid in volume $d\tau$ in one second

or $$\nabla\cdot\rho\mathbf{v} + \frac{d\rho}{dt} = 0 \tag{2.48}$$

Equation (2.47) is called the *equation of continuity* and it manifests the principle of conservation of mass of the fluid.

If the fluid is incompressible, its density will not change, i.e. $\rho =$ constant, and we get equation (2.46).

However, there may exist some sources or sinks of the fluid in volume element $d\tau$. In that case, we need to modify equation (2.47). Let $\psi(x, y, z)$ be the *net* source density. Here, ψ represents the net rate of creation of mass of fluid per unit volume per unit time. This creation of fluid must not be confused with the creation or destruction of matter. What is meant by creation in the present case is that we are considering a system of flow of fluid of a fixed amount to which some fluid from outside is added. Then equation (2.46) will have to be written as follows:

$$\left.\begin{matrix}\text{(rate of outflow of fluid)} \\ -\text{(rate of creation of fluid)}\end{matrix}\right\} = \text{(rate of decrease of fluid)} \qquad (2.49)$$

or for volume $d\tau$,

$$\nabla \cdot \mathbf{V}\, d\tau - \psi\, d\tau = -\frac{d\rho}{dt}\, d\tau$$

This can be put into the usual form

$$\nabla \cdot \mathbf{V} + \frac{d\rho}{dt} = \psi \qquad (2.50)$$

When $\psi = 0$, equation (2.50) reduces to the equation of continuity (2.47).

2.8 CURL OF A VECTOR POINT FUNCTION

We have already remarked that when a vector field is expressible as the gradient of a scalar field, the line integral of the vector along a closed path is zero. This result is independent of the actual path followed. Such a vector field was named a lamellar field. There is, however, a large number of vector fields in physics in which this property is not observed.

Consider a small region of such a vector field (Fig. 2.10). In this region, for the sake of simplicity, assume that the given vector $\mathbf{V}$ has the same direction at all points but different magnitude at various points over this volume. Let us consider, for convenience, rectangular path $ABCD$ such that the plane of the rectangle is perpendicular to the direction of vector $\mathbf{V}$. Then line integral $\oint \mathbf{V} \cdot d\mathbf{r}$ along this path is zero since $\mathbf{V}$ is perpendicular to $d\mathbf{r}$ everywhere.

Now, let us suppose that the plane of the rectangular area is made parallel to the direction of vector $\mathbf{V}$. Then $\oint \mathbf{V} \cdot d\mathbf{r}$ over the closed path $A'B'C'D'$ will have a finite value. This is because the value of the vector along $A'B'$ is different from that along $C'D'$. In a similar manner we can argue that the line integral has finite values for intermediate orientations. The magnitude of this line integral will depend upon orientation of the rectangle with respect to vector $\mathbf{V}$.

The maximum value of this line integral expressed per unit area which is enclosed by the path of integration is called the *curl* of the vector field at the point. The curl of vector **V** is a vector drawn in the direction of the positive normal to the area under consideration when in the position in which the maximum value of the line integral is obtained.

In some books the word rotation (briefly rot) is used for curl.

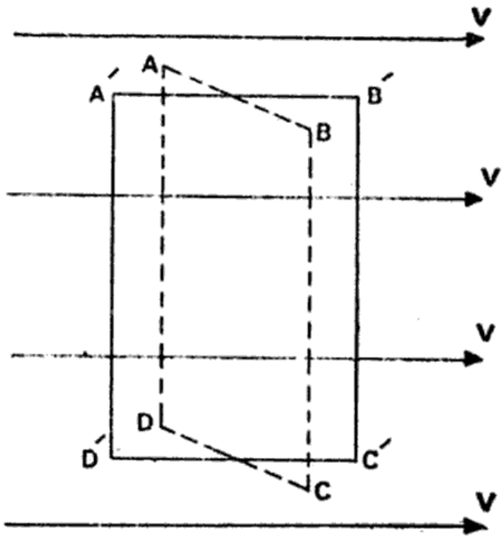


Fig. 2.10 Two positions of rectangular paths in a vector field **V**; $A'B'C'D'$ is parallel and $ABCD$ is perpendicular to **V** everywhere

As before let **V** be the vector point function having components V_x, V_y and V_z along the x, y, z axes respectively. Let $\frac{\partial V_x}{\partial y}, \frac{\partial V_y}{\partial x}, \ldots$ represent the rates of change in V_x, V_y, ... with respect to y, x, ... respectively.

Consider the rectangular path which encloses surface marked 1 perpendicular to the y component of **V** (Fig. 2.11). The values of the x and

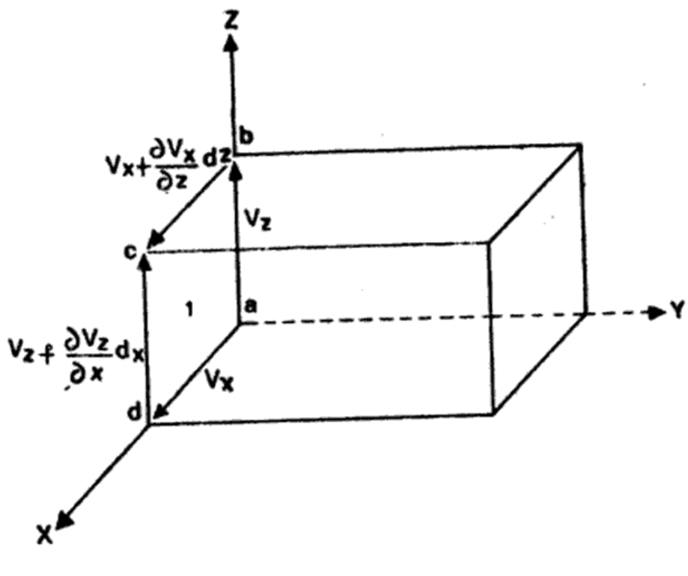


Fig. 2.11 Rectangular path *abcd* in plane $y = 0$ in a field with vector **V** at the origin

z components of the vector **V** along *ad*, *bc*, *ab* and *dc* are V_x, $V_x + \frac{\partial V_x}{\partial z} dz$, V_z and $V_z + \frac{\partial V_z}{\partial x} dx$ respectively.

Hence, the line integral around the contour *abcda* is

$$V_z\, dz + \left(V_x + \frac{\partial V_x}{\partial z} dz\right) dx - \left(V_z + \frac{\partial V_z}{\partial x} dx\right) dz - V_x\, dx$$

$$= \left(\frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}\right) dx\, dz$$

Note that the contributions towards the line integral along *cd* and *da* are negative since these paths are described in a sense opposite to the vectors concerned. Further, this line integral is maximum since the paths *ab*, *bc*, *cd* and *da* are taken either parallel or antiparallel to the vectors. Since the value of this maximum line integral per unit area is a vector along the positive normal to the area, we can write

$$\text{curl}_y\, \mathbf{V} = \left(\frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}\right) \tag{2.51}$$

Similarly, we can consider rectangular paths whose planes are perpendicular to the x and z axes and show that

$$\text{curl}_x\, \mathbf{V} = \left(\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}\right) \tag{2.52}$$

and

$$\text{curl}_z\, \mathbf{V} = \left(\frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y}\right) \tag{2.53}$$

Adding these components vectorially, we get

$$\text{curl}\, \mathbf{V} = \mathbf{i}\, \text{curl}_x\, \mathbf{V} + \mathbf{j}\, \text{curl}_y\, \mathbf{V} + \mathbf{k}\, \text{curl}_z\, \mathbf{V}$$

or

$$\text{curl}\, \mathbf{V} = \mathbf{i}\left(\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}\right) + \mathbf{j}\left(\frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}\right) + \mathbf{k}\left(\frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y}\right). \tag{2.54}$$

Equation (2.54) can be written as

$$\text{curl}\, \mathbf{V} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ V_x & V_y & V_z \end{vmatrix} \tag{2.55}$$

In terms of vector differential operator ∇, we can write

$$\text{curl}\, \mathbf{V} = \nabla \times \mathbf{V} \tag{2.56}$$

If curl $\mathbf{V} = 0$, vector **V** is called an *irrotational* vector. For example, it can be shown that the gravitational and electrostatic fields are irrotational fields.

2.9 MORE ABOUT THE VECTOR DIFFERENTIAL OPERATOR ∇

It must be remembered that ∇ is a vector differential operator and we can operate it on a vector point function to give the scalar function divergence or the vector function curl. It can also be dotted with vector **V**

by retaining its operator nature. Thus,

$$\mathbf{V}\cdot\nabla = V_x \frac{\partial}{\partial x} + V_y \frac{\partial}{\partial y} + V_z \frac{\partial}{\partial z} \tag{2.57}$$

and is a scalar operator. Scalar operator of equation (2.57) can operate on a scalar or a vector. But, $\nabla\cdot\mathbf{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}$ represents divergence of vector $\mathbf{V}$ and is a scalar quantity. Further,

$$\mathbf{V}\times\nabla = \mathbf{i}\left(V_y \frac{\partial}{\partial z} - V_z \frac{\partial}{\partial y}\right) + \mathbf{j}\left(V_z \frac{\partial}{\partial x} - V_x \frac{\partial}{\partial z}\right) + \mathbf{k}\left(V_x \frac{\partial}{\partial y} - V_y \frac{\partial}{\partial x}\right) \tag{2.58}$$

or in general,

$$(\mathbf{V}\times\nabla)_i = V_j \frac{\partial}{\partial x_k} - V_k \frac{\partial}{\partial x_j} \tag{2.59}$$

where i, j and k should be given values 1, 2 and 3 in cyclic order. (Further $x_1 = x$, $x_2 = y$ and $x_3 = z$. Similarly, $V_1 = V_x$, $V_2 = V_y$ and $V_3 = V_z$.) Equation (2.58) gives us a vector differential operator. It can operate on a scalar or a vector function.

The nature of the vector differential operator must always be remembered. Thus, $\nabla\times\mathbf{V}$ is a vector, viz. the curl $\mathbf{V}$ while $\mathbf{V}\times\nabla$ is a vector differential operator.

Consider a case when ∇ is to be operated on a product of two functions. Then it must operate on both functions forming the right type of products. Thus,

$$\nabla\times\mathbf{A}f(\mathbf{r}) = \nabla_A\times\mathbf{A}f(\mathbf{r}) + \nabla_f\times\mathbf{A}f(\mathbf{r}) \tag{2.60}$$

The symbol ∇_A indicates ~~that~~ ∇ operates only on vector $\mathbf{A}$, while ∇_f indicates that ∇ operates only on $f(\mathbf{r})$.

Equation (2.60) can then be written as

$$\begin{aligned}\nabla\times\mathbf{A}f(\mathbf{r}) &= (\nabla\times\mathbf{A})f(\mathbf{r}) + [\nabla f(\mathbf{r})]\times\mathbf{A}\\ &= f(\mathbf{r})\nabla\times\mathbf{A} + \nabla f(\mathbf{r})\times\mathbf{A}\end{aligned} \tag{2.61}$$

In equation (2.61), we have dropped the brackets since the terms are unambiguous.

Consider the operation represented by $\nabla\times(\mathbf{A}\times\mathbf{B})$. We can expand it to

$$\nabla\times(\mathbf{A}\times\mathbf{B}) = \nabla_A\times(\mathbf{A}\times\mathbf{B}) + \nabla_B\times(\mathbf{A}\times\mathbf{B}) \tag{2.62}$$

Using the expansion of vector triple product, we can write

$$\nabla\times(\mathbf{A}\times\mathbf{B}) = \mathbf{A}(\mathbf{B}\cdot\nabla_A) - \mathbf{B}(\nabla_A\cdot\mathbf{A}) + \mathbf{A}(\nabla_B\cdot\mathbf{B}) - \mathbf{B}(\mathbf{A}\cdot\nabla_B)$$

or $$\nabla\times(\mathbf{A}\times\mathbf{B}) = (\mathbf{B}\cdot\nabla)\mathbf{A} - \mathbf{B}(\nabla\cdot\mathbf{A}) + \mathbf{A}(\nabla\cdot\mathbf{B}) - (\mathbf{A}\cdot\nabla)\mathbf{B} \tag{2.63}$$

In equation (2.63) the suffixes have been dropped. It should be noted that $\mathbf{B}\cdot\nabla$ and $\mathbf{A}\cdot\nabla$ are the operators, they operate on vectors $\mathbf{A}$ and $\mathbf{B}$ respectively. Since, the left-hand side of the above equations is a vector, each term on the right-hand side must also be a vector. Hence, $\mathbf{A}$ and $\mathbf{B}$ in the first and the last terms on the right-hand side are rearranged to get vectors.

2.10 ILLUSTRATION OF CURL OF A VECTOR—ANGULAR VELOCITY

Consider a rigid body rotating about an axis passing through the point O (Fig. 2.12). Let it rotate with a constant angular velocity $\boldsymbol{\omega}$. Any point P in the body such that $\mathbf{OP} = \mathbf{r}$ is moving with a linear velocity $\mathbf{v}$ given by

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r} \tag{2.64}$$

If, in addition, the body as a whole is having a constant translational velocity $\mathbf{v}_0$, the total linear velocity of the particle at P is

$$\mathbf{V} = \mathbf{v}_0 + \boldsymbol{\omega} \times \mathbf{r} \tag{2.65}$$

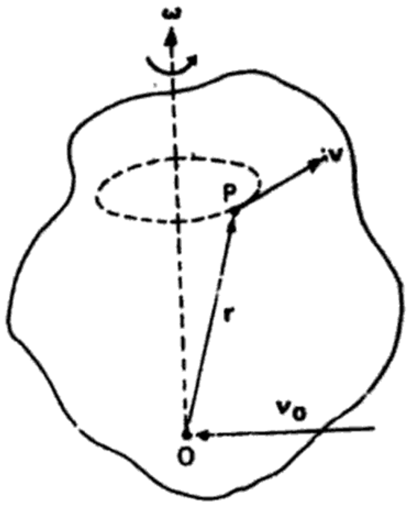


Fig. 2.12 Angular velocity $\boldsymbol{\omega}$ of the rigid body can be expressed as a curl of the linear velocity of the point P in the body

Now, let us calculate curl $\mathbf{V}$. Thus,

$$\nabla \times \mathbf{V} = \nabla \times \mathbf{v}_0 + \nabla \times (\boldsymbol{\omega} \times \mathbf{r})$$

But, $\nabla \times \mathbf{v}_0 = 0$, since $\mathbf{v}_0$ is a constant vector.

Further,

$$\begin{aligned}\nabla \times (\boldsymbol{\omega} \times \mathbf{r}) &= \nabla_\omega \times (\boldsymbol{\omega} \times \mathbf{r}) + \nabla_r \times (\boldsymbol{\omega} \times \mathbf{r}) \\ &= (\mathbf{r} \cdot \nabla_\omega)\boldsymbol{\omega} - \mathbf{r}(\nabla_\omega \cdot \boldsymbol{\omega}) + \boldsymbol{\omega}(\nabla_r \cdot \mathbf{r}) - (\boldsymbol{\omega} \cdot \nabla_r)\mathbf{r}\end{aligned}$$

But $\boldsymbol{\omega}$ is also a constant vector. Hence, any ∇ operation on $\boldsymbol{\omega}$ gives zero result.

Hence, we get,

$$\begin{aligned}\nabla \times \mathbf{V} = \nabla \times (\boldsymbol{\omega} \times \mathbf{r}) &= \boldsymbol{\omega}(\nabla \cdot \mathbf{r}) - (\boldsymbol{\omega} \cdot \nabla)\mathbf{r} \\ &= 3\boldsymbol{\omega} - \boldsymbol{\omega}\end{aligned}$$

since $\nabla \cdot \mathbf{r} = 3$

and
$$\begin{aligned}(\boldsymbol{\omega} \cdot \nabla)\mathbf{r} &= \left(\omega_x \frac{\partial}{\partial x} + \omega_y \frac{\partial}{\partial y} + \omega_z \frac{\partial}{\partial z}\right)(\mathbf{i}x + \mathbf{j}y + \mathbf{k}z) \\ &= \mathbf{i}\omega_x + \mathbf{j}\omega_y + \mathbf{k}\omega_z = \boldsymbol{\omega}\end{aligned}$$

Hence
$$\nabla \times \mathbf{V} = 2\boldsymbol{\omega} \tag{2.66}$$

or
$$\boldsymbol{\omega} = \tfrac{1}{2}\nabla \times \mathbf{V} = \tfrac{1}{2}\,\text{curl}\,\mathbf{V}$$

Thus, the curl of the linear velocity of a particle of a rigid body is equal to twice the angular velocity. Curl $\mathbf{V}$, therefore, describes the rotation of a vector field $\mathbf{V}$ and hence the name *curl* or *rotation* is justified.

In the term $\boldsymbol{\omega} \cdot \nabla\mathbf{r}$ in the above expansion, the product $\nabla\mathbf{r}$ behaves like a unit operator since $\boldsymbol{\omega} \cdot \nabla\mathbf{r} = \boldsymbol{\omega}$. Operator $\nabla\mathbf{r}$ is called a unit dyadic.

2.11 MULTIPLE DEL OPERATIONS

We have already defined the gradient of a scalar point function and the divergence and curl of vector point function. We can further operate the del operator on these to get the following multiple del operations.

$$\left.\begin{array}{llll}\text{(a)} & \text{div grad } \Phi = \nabla\cdot\nabla\Phi & \text{(d)} & \text{div curl } \mathbf{V} = \nabla\cdot(\nabla\times\mathbf{V}) \\ \text{(b)} & \text{curl grad } \Phi = \nabla\times\nabla\Phi & \text{(e)} & \text{curl curl } \mathbf{V} = \nabla\times(\nabla\times\mathbf{V}) \\ \text{(c)} & \text{grad div } \mathbf{V} = \nabla(\nabla\cdot\mathbf{V}) & \text{(f)} & \text{Laplacian } \mathbf{V} = \nabla\cdot\nabla\mathbf{V}\end{array}\right\} \quad (2.67)$$

All the six operations give second-order derivatives which occur in the second-order differential equations in physics, particularly in electrodynamics.

The expressions (a) and (f) involve a second-order differential operator called the *Laplacian*. This is a scalar operator defined by

$$\nabla^2 = \nabla\cdot\nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\cdot \quad (2.68)$$

The Laplacian operator operates on scalar point function Φ as in (a) and also on vector point function $\mathbf{V}$ as in (f) to give a scalar and a vector respectively.

Thus,

$$\text{div grad } \Phi = \nabla\cdot\nabla\Phi = \nabla^2\Phi = \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} \quad (2.69)$$

and

$$\nabla\cdot\nabla\mathbf{V} = \nabla^2\mathbf{V} = \frac{\partial^2\mathbf{V}}{\partial x^2} + \frac{\partial^2\mathbf{V}}{\partial y^2} + \frac{\partial^2\mathbf{V}}{\partial z^2} \quad (2.70)$$

The Laplacian operator occurs in a large number of important equations in theoretical physics. We quote a few equations which involve the Laplacian operator

(i) $\nabla^2\Phi = 0$ — Laplace's equation

(ii) $\nabla^2\Phi = \frac{\rho}{\epsilon_0}$ — Poisson's equation

(iii) $\nabla^2\Phi = \frac{1}{v^2}\frac{\partial^2\Phi}{\partial t^2}$ — Wave equation

(iv) $\nabla^2\Phi = \frac{1}{h^2}\frac{\partial\Phi}{\partial t}$ — Equation ot conduction of heat

(v) $\nabla^2\Phi = \frac{2m}{\hbar^2}(V - E)\Phi$ — Schrödinger's wave equation

Operation curl grad Φ mentioned in (b) is

$$\text{curl grad } \Phi = \nabla\times(\nabla\Phi)$$

$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial\Phi}{\partial x} & \frac{\partial\Phi}{\partial y} & \frac{\partial\Phi}{\partial z} \end{vmatrix}$$

$$= \mathbf{i}\left(\frac{\partial^2\Phi}{\partial y\,\partial z} - \frac{\partial^2\Phi}{\partial z\,\partial y}\right) + \mathbf{j}\left(\frac{\partial^2\Phi}{\partial z\,\partial x} - \frac{\partial^2\Phi}{\partial x\,\partial z}\right) + \mathbf{k}\left(\frac{\partial^2\Phi}{\partial y\,\partial x} - \frac{\partial^2\Phi}{\partial x\,\partial y}\right)$$

or

$$\text{curl (grad } \Phi) = 0 \quad (2.71)$$

In writing equation (2.71), we have assumed that the order of partial differentiation can be interchanged and will be true if second-order partial derivatives of Φ are continuous functions of argument x, y, z.

Operation grad div $\mathbf{V}$ mentioned in (c) is a vector and appears in the expansion of curl curl $\mathbf{V}$ mentioned in (e).

Now,
$$\text{curl curl } \mathbf{V} = \nabla \times \nabla \times \mathbf{V}$$
$$= \nabla(\nabla \cdot \mathbf{V}) - \nabla \cdot \nabla \mathbf{V} \tag{2.72}$$
or
$$\text{curl curl } \mathbf{V} = \text{grad div } \mathbf{V} - \nabla^2 \mathbf{V}$$

Operation grad div $\mathbf{V}$ can be written as

$$\nabla(\nabla \cdot \mathbf{V}) = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right)\left(\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}\right)$$
$$= \mathbf{i}\left(\frac{\partial^2 V_x}{\partial x^2} + \frac{\partial^2 V_y}{\partial x\, \partial y} + \frac{\partial^2 V_z}{\partial x\, \partial z}\right) + \mathbf{j}\left(\frac{\partial^2 V_x}{\partial y\, \partial x} + \frac{\partial^2 V_y}{\partial y^2} + \frac{\partial^2 V_z}{\partial y\, \partial z}\right)$$
$$+ \mathbf{k}\left(\frac{\partial^2 V_x}{\partial z\, \partial x} + \frac{\partial^2 V_y}{\partial z\, \partial y} + \frac{\partial^2 V_z}{\partial z^2}\right) \tag{2.73}$$

In expression (f), viz. $\nabla \cdot \nabla \mathbf{V}$, $\nabla \mathbf{V}$ is a dyadic and divergence of a dyadic gives a vector.

Operation div curl $\mathbf{V}$ is a scalar and can be expanded as a scalar triple product. Thus,

$$\nabla \cdot \nabla \times \mathbf{V} = \begin{vmatrix} \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ V_x & V_y & V_z \end{vmatrix}$$
or
$$\nabla \cdot \nabla \times \mathbf{V} = 0 \tag{2.74}$$

2.12 IRROTATIONAL AND SOLENOIDAL VECTORS

We have already remarked that vector $\mathbf{V}$ is solenoidal, if
$$\text{div } \mathbf{V} = \nabla \cdot \mathbf{V} = 0$$

But, the divergence of the curl of a vector is zero, i.e., $\nabla \cdot \nabla \times \mathbf{A} = 0$ in accordance with equation (2.74). Hence, solenoidal vector $\mathbf{V}$ can always be expressed as a curl of a vector function. This latter vector function is called a *vector potential*.

For example, for magnetic induction $\mathbf{B}$, $\nabla \cdot \mathbf{B} = 0$. Hence,
$$\mathbf{B} = \nabla \times \mathbf{A} \tag{2.75}$$
where $\mathbf{A}$ is called the magnetic vector potential. Vector potential $\mathbf{A}$ is not unique. Suppose we add another vector grad Φ to $\mathbf{A}$, we get
$$\mathbf{B}' = \nabla \times [\mathbf{A} + \text{grad } \Phi] = \nabla \times \mathbf{A} + \nabla \times \nabla\Phi \tag{2.76}$$
But, $\nabla \times \nabla\Phi = 0$ by virtue of equation (2.71). Hence, we still get $\mathbf{B} = \nabla \times \mathbf{A}$.

We have also remarked that vector $\mathbf{V}$ is irrotational if curl $\mathbf{V} = \nabla \times \mathbf{V} = 0$. But, curl of gradient of a scalar is always zero, i.e., $\nabla \times \nabla\Phi = 0$ by

equation (2.71). Hence, an irrotational vector $\mathbf{V}$ can always be expressed as gradient of some scalar function. Thus, $\mathbf{V} = \nabla\Phi = \text{grad}\ \Phi$. This scalar function Φ is called a *scalar potential.* Thus, in the case of gravitational or electrostatic field, we have

$$\nabla \times \mathbf{E} = 0 \quad \text{or} \quad \mathbf{E} = -\nabla\Phi \tag{2.77}$$

The negative sign indicates that vector $\mathbf{E}$ is directed in a direction along which Φ decreases with increases $d\mathbf{r}$ in $\mathbf{r}$. The potential Φ is also not unique because addition of a scalar constant function c to Φ would also yield

$$-\nabla(\Phi + c) = -\nabla\Phi = \mathbf{E}$$

We now illustrate the use of multiple del operations by two examples.

Example 1. Find $\nabla \cdot \nabla f(r)$, where $r = \sqrt{x^2 + y^2 + z^2}$

We can write

$$\nabla f(r) = \mathbf{i}\frac{\partial f(r)}{\partial x} + \mathbf{j}\frac{\partial f(r)}{\partial y} + \mathbf{k}\frac{\partial f(r)}{\partial z}$$

But,

$$\frac{\partial f(r)}{\partial x} = \frac{\partial f(r)}{\partial r}\cdot\frac{\partial r}{\partial x} = \frac{x}{r}\frac{\partial f(r)}{\partial x}, \text{ since } \frac{\partial r}{\partial x} = \frac{x}{r}$$

Hence,

$$\nabla f(r) = \mathbf{i}\frac{x}{r}\frac{\partial f(r)}{\partial r} + \mathbf{j}\frac{y}{r}\frac{\partial f(r)}{\partial r} + \mathbf{k}\frac{z}{r}\cdot\frac{\partial f(r)}{\partial r}$$

$$= \frac{\mathbf{r}}{r}\frac{\partial f(r)}{\partial r}$$

Hence,

$$\nabla \cdot \nabla f(r) = \nabla \cdot \frac{\mathbf{r}}{r}\frac{\partial f(r)}{\partial r}$$

$$= \frac{\partial}{\partial x}\left[\frac{x}{r}\frac{\partial f(r)}{\partial r}\right] + \frac{\partial}{\partial y}\left[\frac{y}{r}\frac{\partial f(r)}{\partial r}\right] + \frac{\partial}{\partial z}\left[\frac{z}{r}\frac{\partial f(r)}{\partial r}\right]$$

Let us now evaluate the first term on the right-hand side

$$\frac{\partial}{\partial x}\left[\frac{x}{r}\frac{\partial f(r)}{\partial r}\right] = \frac{1}{r}\frac{\partial f(r)}{\partial r} - \frac{x}{r^2}\frac{\partial r}{\partial x}\frac{\partial f(r)}{\partial r} + \frac{x}{r}\frac{\partial^2 f(r)}{\partial r^2}\frac{\partial r}{\partial x}$$

$$= \frac{1}{r}\frac{\partial f(r)}{\partial r} - \frac{x^2}{r^3}\frac{\partial f(r)}{\partial r} + \frac{x^2}{r^2}\frac{\partial^2 f(r)}{\partial r^2}$$

Similarly we can evaluate the other terms or write them by noting the regularity of the appearance of variables x, y and z. Adding all these three terms, we get

$$\nabla \cdot \nabla f(r) = \frac{3}{r}\frac{\partial f(r)}{\partial r} - \frac{(x^2 + y^2 + z^2)}{r^3}\frac{\partial f(r)}{\partial r} + \frac{(x^2 + y^2 + z^2)}{r^2}\frac{\partial^2 f(r)}{\partial r^2}$$

$$= \frac{2}{r}\frac{\partial f(r)}{\partial r} + \frac{\partial^2 f(r)}{\partial r^2}$$

If $f(r) = r^n$, then,

$$\nabla \cdot \nabla(r^n) = \nabla^2 r^n = \frac{2}{r}\, nr^{n-1} + n(n-1)r^{n-2}$$

$$= n(n+1)r^{n-2}$$

If $n = 0$, i.e. if $f(r) = \text{constant}$,

$$\nabla \cdot \nabla f(r) = 0$$

If $n = -1$, i.e. if $f(r) = \frac{1}{r}$, and $r \neq 0$

$$\nabla \cdot \nabla f(r) = \nabla^2 \left(\frac{1}{r}\right) = 0$$

But, $\nabla^4 f(r) = 0$ is Laplace's equation. Hence, $f(r) = \frac{1}{r}$ is a solution of Laplace's equation.

Example 2: Given $\mathbf{V} = \mathbf{i}(x^2 + 2yz) + \mathbf{j}\, 2yz - \mathbf{k}(z^2 + 2zx)$, find a vector $\mathbf{A}$ such that $\nabla \times \mathbf{A} = \mathbf{V}$.

Solution: We first find the divergence of $\mathbf{V}$:

$$\text{Div } \mathbf{V} = \frac{\partial}{\partial x}(x^2 + 2yz) + \frac{\partial}{\partial y}(2yz) + \frac{\partial}{\partial z}\{-(z^2 + 2zx)\}$$

$$= 2x + 2z - 2z - 2x = 0.$$

This shows that vector $\mathbf{V}$ must be solenoidal. Vector $\mathbf{A}$ cannot be determined uniquely and hence to find its y and z components, let us assume that $A_x = 0$, as a particular case. Then, by using curl $\mathbf{A} = \mathbf{V}$,

$$\text{curl}_y\, \mathbf{A} = -\frac{\partial A_z}{\partial x} = 2yz$$

and

$$\text{curl}_z\, \mathbf{A} = \frac{\partial A_y}{\partial x} = -(z^2 + 2zx)$$

On integration, we get

$$A_z = -2yzx + C_1(y, z)$$

and

$$A_y = -(z^2x + zx^2) + C_2(y, z)$$

where C_1 and C_2 are the constants of integration which depend upon y and z.

Now, x component of $\mathbf{V}$ is given by,

$$V_x = x^2 + 2yz$$

$$= \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}$$

Hence,

$$V_x = -2zx + \frac{\partial C_1}{\partial y} + 2zx + x^2 - \frac{\partial C_2}{\partial z}$$

Equating the values of V_x, we get

$$x^2 + 2yz = \frac{\partial C_1}{\partial y} + x^2 - \frac{\partial C_2}{\partial z}$$

This equality will be satisfied only if we choose

$C_1 = 0,\ C_2 = -yz^2$ or $C_1 = y^2z,\ C_2 = 0$ or $C_1 = \frac{1}{2}y^2z,\ C_2 = \frac{1}{2}yz^2$

Hence, with one of the choice vector $\mathbf{A}$ can be expressed as

$$\mathbf{A} = -\mathbf{j}(z^2x + zx^2) - \mathbf{k}(2xyz - y^2z)$$

2.13 SOME USEFUL IDENTITIES

Following are some identities that can be proved easily:

1. grad $(\Phi\Psi) = \Phi$ grad $\Psi + \Psi$ grad Φ.
2. div $(\Phi\mathbf{A}) =$ grad $\Phi\cdot\mathbf{A} + \Phi$ div $\mathbf{A}$.
3. curl $(\Phi\mathbf{A}) =$ grad $\Phi\times\mathbf{A} + \Phi$ curl $\mathbf{A}$.
4. grad $(\mathbf{A}\cdot\mathbf{B}) = (\mathbf{A}\cdot\text{grad})\,\mathbf{B} + (\mathbf{B}\cdot\text{grad})\,\mathbf{A} + \mathbf{A}\times\text{curl}\,\mathbf{B} + \mathbf{B}\times\text{curl}\,\mathbf{A}$.
5. div $(\mathbf{A}\times\mathbf{B}) = \mathbf{B}\cdot\text{curl}\,\mathbf{A} - \mathbf{A}\cdot\text{curl}\,\mathbf{B}$.
6. curl $(\mathbf{A}\times\mathbf{B}) = \mathbf{A}\,(\text{div}\,\mathbf{B}) - \mathbf{B}\,(\text{div}\,\mathbf{A}) + (\mathbf{B}\cdot\text{grad})\,\mathbf{A} - (\mathbf{A}\cdot\text{grad})\,\mathbf{B}$.
7. curl grad $\Phi = \nabla\times\nabla\Phi = 0$.
8. div curl $\mathbf{A} = \nabla\cdot\nabla\times\mathbf{A} = 0$.
9. curl curl $\mathbf{A} = \nabla\times\nabla\times\mathbf{A} = \nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A}$
   $=$ grad div $\mathbf{A}$ $-$ Laplacian $\mathbf{A}$.
10. div (grad $\Phi\times$ grad Ψ) $= \nabla\cdot(\nabla\Phi\times\nabla\Psi) = 0$.

2.14 GAUSS' THEOREM

Gauss' divergence theorem is one of the most important theorems of vector analysis and it relates a volume integral with the surface integral. It states that *the volume integral of the divergence of vector point function* $\mathbf{V}$ *taken over any volume* τ *is equal to the surface integral of vector* $\mathbf{V}$ *taken over the closed surface surrounding volume* τ.

Thus,

$$\int_\tau \nabla\cdot\mathbf{V}\,d\tau = \int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma} \tag{2.78}$$

where $d\tau = dx\,dy\,dz$, a volume element and $d\boldsymbol{\sigma}$ is a vector element of area.

Proof: Consider a large volume τ and imagine it to be subdivided into volume elements $d\tau_i$ (Fig. 2.13). The outward flux of vector $\mathbf{V}$ from $d\tau_i$ or the outflow from each element is $\nabla\cdot\mathbf{V}\,d\tau_i$. Hence, the total

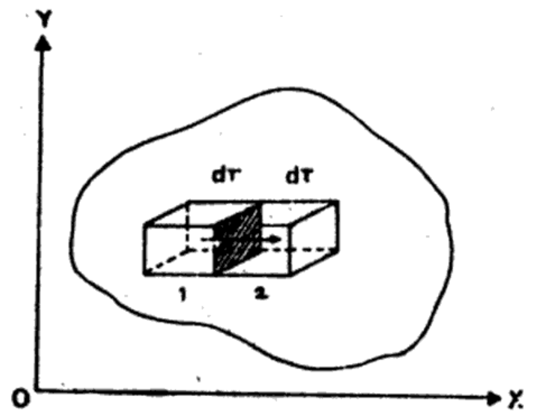


Fig. 2.13 Two volume elements in contact. The outflows across the interior surfaces cancel each other

outflow from all the volume elements is

$$\sum_i \nabla \cdot \mathbf{V}\, d\tau_i \tag{2.79}$$

Consider the two neighbouring volume elements marked 1 and 2. The outflow from element 1 across the common surface is the inflow for element 2. Hence, in taking the sum, these factors cancel each other. Therefore, the outflows across interior surfaces cancel each other and the total outflow from all the elements in the entire volume τ is then equal to the outflow across the closed surface surrounding volume τ. As volume element $d\tau_i \to 0$, the total outflow can be expressed as

$$\int_\tau \nabla \cdot \mathbf{V}\, d\tau \text{ and } \int_\sigma \mathbf{V} \cdot \mathbf{n}\, d\sigma \text{ or } \int_\sigma \mathbf{V} \cdot d\boldsymbol{\sigma} \text{ respectively}$$

Hence, we get

$$\int_\tau \nabla \cdot \mathbf{V}\, d\tau = \int_\sigma \mathbf{V} \cdot d\boldsymbol{\sigma} \tag{2.80}$$

A more rigorous proof can also be given:

We write $\int_\tau \nabla \cdot \mathbf{V}\, d\tau$ as

$$\int_\tau \nabla \cdot \mathbf{V}\, d\tau = \iiint \frac{\partial V_x}{\partial x}\, dx\, dy\, dz + \iiint \frac{\partial V_y}{\partial y}\, dx\, dy\, dz + \iiint \frac{\partial V_z}{\partial z}\, dx\, dy\, dz \tag{2.81}$$

Consider the first term on the right-hand side of equation (2.81).

Integrating with respect to x, i.e., along a strip of cross-section $dy\, dz$ extending from point P_1 to P_2 (Fig. 2.14) which are on the bounding surface, we get

$$\iiint \frac{\partial V_x}{\partial x}\, dx\, dy\, dz = \iint \{V_x(x_2, y, z) - V_x(x_1, y, z)\}\, dy\, dz \tag{2.82}$$

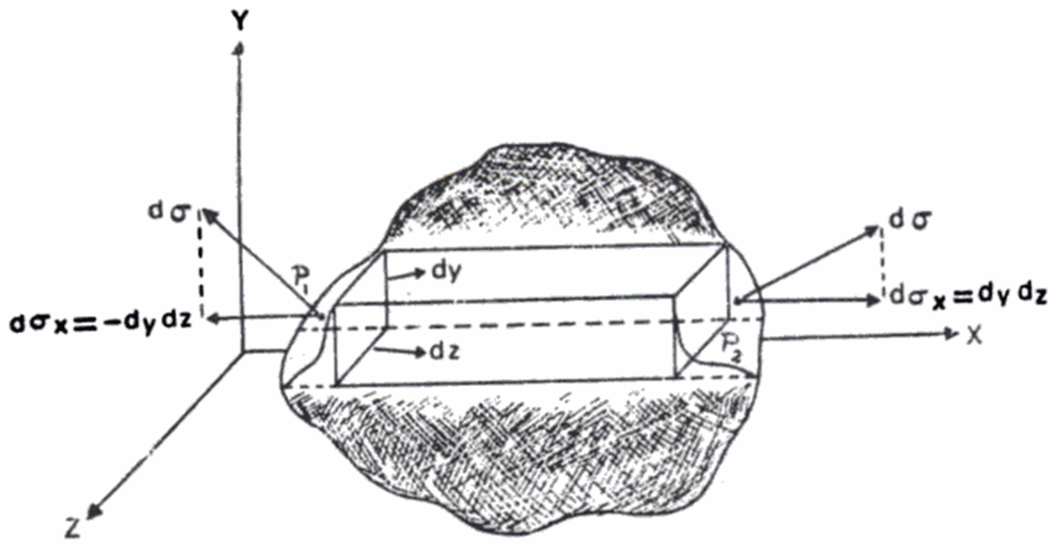


Fig. 2.14 A strip of volume parallel to the x-axis cutting the enveloping surface

where (x_1, y, z) and (x_2, y, z) are the coordinates of points P_1 and P_2 respectively.

Now, at point P_1,

$$dy\, dz = -\, d\sigma_x$$

and at point P_2,

$$dy\, dz = +\, d\sigma_x$$

Substituting these values in equation (2.82), we get

$$\iiint \frac{\partial V_x}{\partial x}\, dx\, dy\, dz = \int_\sigma V_x(x_2, y, z)\, d\sigma_x + \int_\sigma V_x(x_1, y, z)\, d\sigma_x$$

Note that the first surface integral is taken over the right-hand projection part of surface σ whereas the second is taken over the left-hand part of σ.

Hence, in general,

$$\iiint \frac{\partial V_x}{\partial x}\, dx\, dy\, dz = \int_\sigma V_x\, d\sigma_x$$

Similarly,

$$\iiint \frac{\partial V_y}{\partial y}\, dx\, dy\, dz = \int_\sigma V_y\, d\sigma_y$$

and

$$\iiint \frac{\partial V_z}{\partial z}\, dx\, dy\, dz = \int_\sigma V_z\, d\sigma_z$$

Adding the three terms, we get

$$\left.\begin{aligned} &\int_\tau \nabla\cdot\mathbf{V}\, d\tau = \int_\sigma V_x\, d\sigma_x + \int_\sigma V_y\, d\sigma_y + \int_\sigma V_z\, d\sigma_z \\ \text{or}\quad &\int_\tau \nabla\cdot\mathbf{V}\, d\tau = \int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma} \end{aligned}\right\} \tag{2.83}$$

This proves the theorem.

2.15 GREEN'S THEOREM

Gauss' theorem can be put in two different forms. Let $\mathbf{V} = u\nabla v$, where u and v are continuous scalar functions having continuous derivatives. Then,

$$\nabla\cdot(u\nabla v) = \nabla u\cdot\nabla v + u\nabla^2 v \tag{2.84}$$

Similarly, on interchanging u and v, we get

$$\nabla\cdot(v\nabla u) = \nabla v\cdot\nabla u + v\nabla^2 u. \tag{2.85}$$

Then, integrating equation (2.84) over volume τ, we get

$$\int_\tau \nabla\cdot(u\nabla v)\, d\tau = \int_\tau (\nabla u\cdot\nabla v + u\nabla^2 v)\, d\tau$$

Converting the left-hand side into a surface integral by Gauss' theorem, we get

$$\int_\sigma (u\nabla v)\cdot d\boldsymbol{\sigma} = \int_\tau (\nabla u\cdot\nabla v + u\nabla^2 v)\, d\tau \tag{2.86}$$

A similar process applied to equation (2.85) yields

$$\int_\sigma (v\,\nabla u)\cdot d\boldsymbol{\sigma} = \int_\tau (\nabla v\cdot\nabla u + v\,\nabla^2 u)\,d\tau \tag{2.87}$$

Equations (2.86) and (2.87) are known as the *first form of Green's theorem.*

Subtracting equation (2.85) from equation (2.84), we get

$$\nabla\cdot(u\,\nabla v - v\,\nabla u) = u\,\nabla^2 v - v\,\nabla^2 u$$

Integrating this expression over volume τ and converting the left-hand side into a surface integral by Gauss' theorem, we get

$$\int_\sigma (u\,\nabla v - v\,\nabla u)\cdot d\boldsymbol{\sigma} = \int_\tau (u\,\nabla^2 v - v\,\nabla^2 u)\,d\tau \tag{2.88}$$

Equation (2.88) is known as the *second or symmetric form of Green's theorem.*

2.16 VOLUME INTEGRALS OF OTHER TYPES

In Gauss' theorem, we come across a volume integral in which divergence of a vector function is involved. This is the most important form of the divergence theorem. However, the forms in which the gradient of a scalar function and curl of a vector function appear can also be obtained.

Let
$$\mathbf{V}(\mathbf{r}) = V(\mathbf{r})\mathbf{c} \tag{2.89}$$

where $V(\mathbf{r}) \equiv V(x, y, z)$ is the scalar magnitude of vector point function $\mathbf{V}$ and $\mathbf{c}$ is a unit vector along $\mathbf{V}$ and is assumed to be constant. Hence,

$$\nabla\cdot\mathbf{V} = \mathbf{c}\cdot\nabla V \tag{2.90}$$

With this substitution, Gauss' theorem becomes

$$\int_\tau (\mathbf{c}\cdot\nabla V)\,d\tau = \int_\sigma V\mathbf{c}\cdot d\boldsymbol{\sigma} \tag{2.91}$$

But $\mathbf{c}$ is a constant vector and it can be taken out of integration sign. Then, on rearranging the terms, equation (2.91) becomes

$$\mathbf{c}\cdot\left[\int_\tau \nabla V\,d\tau - \int_\sigma V\,d\boldsymbol{\sigma}\right] = 0$$

As $|\mathbf{c}| \neq 0$ and it is arbitrary, we can take it in any direction other than normal to $\left[\int_\tau \nabla V\,d\tau - \int_\sigma V\,d\boldsymbol{\sigma}\right]$. Hence, we get

$$\int_\tau \nabla V\,d\tau = \int_\sigma V\,d\boldsymbol{\sigma} \tag{2.92}$$

Equation (2.92) gives a form of Gauss's theorem involving gradient.

Similarly, let $\mathbf{V} = \mathbf{C}\times\mathbf{A}$, where $\mathbf{C}$ is once again an arbitrary constant vector. Then,

$$\nabla\cdot\mathbf{V} = \nabla\cdot(\mathbf{C}\times\mathbf{A}) = -\mathbf{C}\cdot(\nabla\times\mathbf{A}) \tag{2.93}$$

Substituting this in equation (2.78), viz. $\int_\tau \nabla\cdot\mathbf{V}\,d\tau = \int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma}$, we get

$$-\mathbf{C}\cdot\int_\tau \nabla\times\mathbf{A}\,d\tau = \int_\sigma \mathbf{C}\times\mathbf{A}\cdot d\boldsymbol{\sigma}$$

i.e.,
$$-\mathbf{C}\cdot\int_\tau \nabla\times\mathbf{A}\,d\tau = -\mathbf{C}\cdot\int_\sigma d\boldsymbol{\sigma}\times\mathbf{A}$$

Since **C** is a constant arbitrary vector, we can write

$$\int_\tau \nabla\times\mathbf{A}\,d\tau = \int_\sigma d\boldsymbol{\sigma}\times\mathbf{A} \tag{2.94}$$

One interesting consequence of equations (2.92), (2.78) and (2.94) is that these can be used to define the divergence, the gradient and the curl. Thus, we can define

$$\text{grad}\,\Phi = \nabla\Phi = \lim_{\int d\tau\to 0}\frac{\int_\sigma \Phi\,d\boldsymbol{\sigma}}{\int d\tau} \tag{2.95}$$

$$\text{div}\,\mathbf{V} = \nabla\cdot\mathbf{V} = \lim_{\int d\tau\to 0}\frac{\int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma}}{\int d\tau} \tag{2.96}$$

and
$$\text{curl}\,\mathbf{V} = \nabla\times\mathbf{V} = \lim_{\int d\tau\to 0}\frac{\int_\sigma d\boldsymbol{\sigma}\times\mathbf{A}}{\int d\tau} \tag{2.97}$$

In all these equations, $d\boldsymbol{\sigma}$ represents a vector element of area while $d\tau$ represents a small volume element. Further, σ is the surface that encloses τ.

2.17 STOKES' THEOREM

Stokes' theorem gives the relationship between a surface integral and a line integral. It states that *the surface integral of the curl of vector point function* **V** *taken over any surface* σ *is equal to the line integral of vector point function* **V** *around the periphery of the surface.* Thus,

$$\int_\sigma (\nabla\times\mathbf{V})\cdot d\boldsymbol{\sigma} = \oint \mathbf{V}\cdot d\mathbf{l} \tag{2.98}$$

where $d\mathbf{l}$ is a vector element of the periphery taken in the sense of rotation of a right-hand screw, the tip of which advances in the direction of positive normal to the surface.

Thus, if the integrand is in the form of the curl of vector **V**, the surface integral depends entirely upon the values of vector **V** at the points on the periphery of the surface and not at the points inside. Thus, all surfaces having the same periphery will give the same value of the integral.

Proof: Consider vector field **V** and draw any open surface σ in it such

that the given closed curve is the boundary of the surface. The actual shape of the surface does not matter (Fig. 2.15). The line integral of

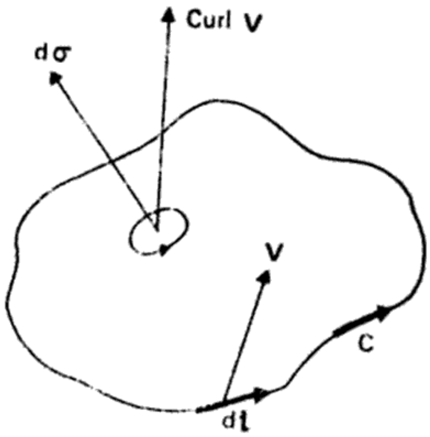


Fig. 2.15 Arbitrary surface in the field of vector **V** bounded by a closed curve C

vector **V** round the closed path when it is traced in the anticlockwise sense as shown in Fig. 2.15 will be denoted by $\oint \mathbf{V}\cdot d\mathbf{l}$.

Consider a small element (shown in Fig. 2.15) of area $d\sigma$. The sense of area vector $d\boldsymbol{\sigma}$ is given by right-hand screw rule.

We want to evaluate a surface integral on the left-hand side of equation (2.98). It is the surface integral of the curl of a vector.

Let us expand and regroup the left-hand side of equation (2.98),

$$\int_\sigma (\nabla\times\mathbf{V})\cdot d\boldsymbol{\sigma} = \int_\sigma\left(\frac{\partial V_x}{\partial z}\,d\sigma_y - \frac{\partial V_x}{\partial y}\,d\sigma_z\right) + \int_\sigma\left(\frac{\partial V_y}{\partial x}\,d\sigma_z - \frac{\partial V_y}{\partial z}\,d\sigma_x\right)$$
$$+ \int_\sigma\left(\frac{\partial V_z}{\partial y}\,d\sigma_x - \frac{\partial V_z}{\partial x}\,d\sigma_y\right) \tag{2.99}$$

To evaluate the right-hand side of equation (2.99), let us orient the axes in such a way that given surface σ intersects plane $x = c$ along some

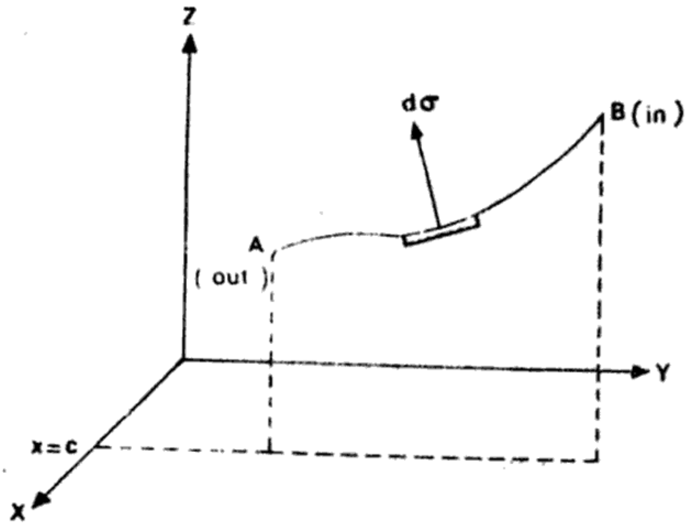


Fig. 2.16 Intersection of the given surface by plane $x = c$

curve AB (Fig. 2.16). The perimeter will appear to come *out* at point A and go *in* at point B. A small elementary area $d\sigma$ is represented by means of vector $d\boldsymbol{\sigma}$ in accordance with the usual convention.

Consider a small strip of the surface between planes $x = c$ and $c + dc$ (Fig. 2.17). On this strip, consider small rectangular area element $d\boldsymbol{\sigma}$, the components of which are

$$d\sigma_y = -dx\,dz \quad \text{and} \quad d\sigma_z = dx\,dy$$

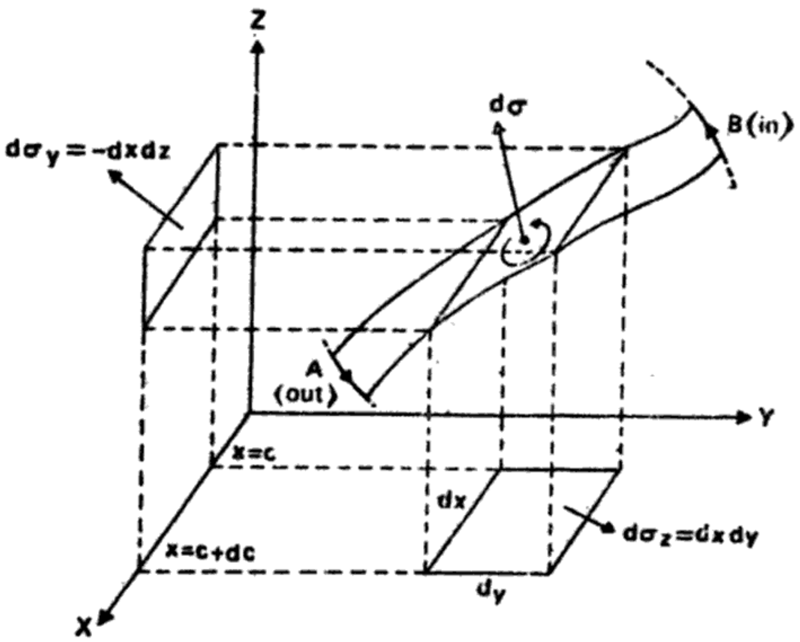


Fig. 2.17 Components of area element $d\boldsymbol{\sigma}$ along $y = 0$ and $z = 0$ planes

The strip is of narrow width and its surface is chosen, for convenience, everywhere to be perpendicular to plane $x = c$. Hence, it does not have x-component. Further, during the process of integration over strip AB, x is fixed. After the integration over this strip, we shall allow x to vary so as to cover the entire surface area drawn with the closed curve as the boundary.

Consider the first term on the right-hand side of equation (2.99). Its integral for area element $d\sigma$ is

$$\left(\frac{\partial V_x}{\partial z}\,d\sigma_y - \frac{\partial V_x}{\partial y}\,d\sigma_z\right) = -\left(\frac{\partial V_x}{\partial z}\,dz + \frac{\partial V_x}{\partial y}\,dy\right)dx$$
$$= -dV_x\,dx$$

In this integration, x is considered to be constant ($x = c$). In the above integral, dV_x is to be evaluated for entire strip AB and hence the first term on the right-hand side of equation (2.99) becomes,

$$\int_\sigma \left(\frac{\partial V_x}{\partial z}\,d\sigma_y - \frac{\partial V_x}{\partial y}\,d\sigma_z\right) = -\int\left[\int_A^B dV_x\right]dx$$
$$= -\int [V_x(x, y_B, z_B) - V_x(x, y_A, z_A)]\,dx$$

where (x, y_B, z_B) and (x, y_A, z_A) are the coordinates of points B and A respectively. This is the integral over the strip of the area ending at points A and B. We have held x fixed and carried out integration over the components of area $d\boldsymbol{\sigma}$ along the y and z axes only. The length of the strip at the periphery at point A is $dx = dl_x$ and that at B is $dx = -dl_x$. Hence, we can write

$$\int_\sigma \left(\frac{\partial V_x}{\partial z}\, d\sigma_y - \frac{\partial V_x}{\partial y}\, d\sigma_z\right) = \int V_x(x, y_B, z_B)\, dl_x + V_x(x, y_A, z_A)\, dl_x$$
$$= \oint V_x\, dl_x \tag{2.100}$$

The last step is written for the entire periphery by allowing the change in x to cover the entire surface.

Similarly we can prove the second and the third terms on the right-hand side of equation (2.99) to be equal to $\oint V_y\, dl_y$ and $\oint V_z\, dl_z$ respectively. Adding all the components, we get

$$\int_\sigma (\nabla \times \mathbf{V})\cdot d\boldsymbol{\sigma} = \oint [V_x\, dl_x + V_y\, dl_y + V_z\, dl_z]$$

or

$$\int_\sigma (\nabla \times \mathbf{V})\cdot d\boldsymbol{\sigma} = \oint \mathbf{V}\cdot d\mathbf{l} \tag{2.101}$$

This proves Stokes' theorem.

As in the case of Gauss' theorem, we can obtain other relations between the surface and the line integrals. These can be stated as

$$\int_\sigma d\boldsymbol{\sigma} \times \nabla\Phi = \oint \Phi d\mathbf{l} \tag{2.102}$$

and

$$\int_\sigma (d\boldsymbol{\sigma} \times \nabla) \times \mathbf{p} = \oint d\mathbf{l} \times \mathbf{p} \tag{2.103}$$

where $\mathbf{V} = \mathbf{c} \times \mathbf{p}$ in which $\mathbf{c}$ is a constant vector.

2.18 PHYSICAL SIGNIFICANCE OF THE CURL OF A VECTOR

Stokes' theorem can be used to define the curl of a vector as follows: Consider element of area $d\boldsymbol{\sigma} = \hat{\mathbf{n}}\, d\sigma$. Then, according to Stokes' theorem for the line integral of vector $\mathbf{V}$ along the boundary of this element of area we must have

$$\oint \mathbf{V}\cdot d\mathbf{l} = (\nabla \times \mathbf{V})\cdot d\boldsymbol{\sigma}$$

or

$$\oint \mathbf{V}\cdot d\mathbf{l} = (\nabla \times \mathbf{V})\cdot \hat{\mathbf{n}}\, d\sigma \tag{2.104}$$

Hence, the component of curl of vector $\mathbf{V}$ along the direction of unit vector $\hat{\mathbf{n}}$ is

$$(\nabla \times \mathbf{V})\cdot \hat{\mathbf{n}} = \lim_{d\sigma \to 0} \frac{1}{d\sigma} \oint \mathbf{V}\cdot d\mathbf{l} \tag{2.105}$$

Line integral $\oint \mathbf{V}\cdot d\mathbf{l}$ around a plane closed curve is also termed the circulation of vector $\mathbf{V}$ through the area enclosed by the curve.

Consider now vector field $\mathbf{V} = \rho\mathbf{v}$, where ρ is the density of liquid and

v its velocity. Consider a small element of area $d\sigma$ in the xy-plane (Fig. 2.18). Imagine a paddle wheel with its axis parallel to the z-axis to be placed in the position of the element of area. This paddle wheel will rotate if **V** is not constant at various points of the wheel. If, however, **V** is constant in magnitude and direction everywhere, the paddle wheel will not rotate. In other words, the line integral or the circulation of vector **V** around $d\sigma$ will be zero.

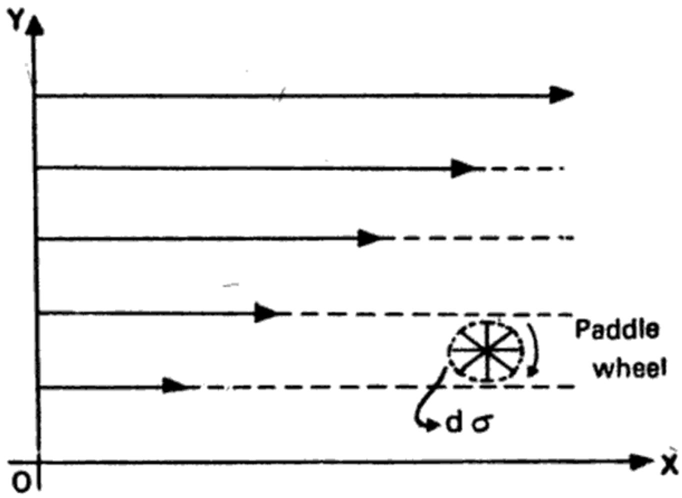


Fig. 2.18 Circulation of vector

Thus, the circulation which is also the component of the curl in the limit as area $d\sigma \to 0$, depends upon the nature of the vector field and also upon the orientation of the area with respect to the vector field. In the above illustration of the flow of a liquid, the circulation will have a maximum value when $d\boldsymbol{\sigma} = \mathbf{k}\, d\sigma$ and zero when the area vector is along x or y direction.

QUESTIONS

1. Can a vector be differentiated with respect to a vector? Explain.
2. Explain the terms: irrotational and solenoidal vector. Give their illustrations.
3. Prove the identity 4 listed in article 2.13.
4. Prove that $\int_\sigma \mathbf{r}\cdot d\boldsymbol{\sigma} = 3\tau$, where τ is the volume enclosed by surface σ. What is the value of τ if the surface of integration is that of a sphere of radius r?
5. State Green's theorem in a plane.
6. Use Green's theorem in a plane to show that the area bounded by a simple closed curve C is $\frac{1}{2}\oint (x\, dy - y\, dx)$.

7. Find the area of ellipse $x = a \cos\theta$ and $y = b \sin\theta$ using the result of question 6.

8. Show that $\int_\sigma \hat{\mathbf{n}}\, d\sigma = 0$ for any closed surface σ.

9. Under what conditions are the divergence of a vector quantity (a) positive, (b) negative, and (c) zero?

10. Show that $\int \mathbf{r} \times d\boldsymbol{\sigma} = 0$ for any closed surface σ.

11. Find the spherical components of velocity and acceleration of a particle at the position given by
$$r = b,\ \theta = \theta_0 \cos \omega t \text{ and } \varphi = \omega t.$$

12. Prove that $\nabla \cdot \hat{\mathbf{r}} = \frac{2}{r}$, where $\hat{\mathbf{r}}$ is a unit vector along $\mathbf{r}$.

13. Prove that $(\mathbf{u}\cdot\nabla)\mathbf{u} = \mathbf{u}$, i.e., $\mathbf{u}\cdot\nabla$ is a unit operator.

14. Determine constant c such that the vector
$$\mathbf{A} = (x + 3y)\mathbf{i} + (y - 2z)\mathbf{j} + (x + cz)\mathbf{k}$$
is solenoidal.

15. $\Phi(x, y, z)$ is a solution of Laplace's equation. Show that $\nabla\Phi$ is both irrotational and solenoidal.

16. A force field is given by
$$\mathbf{F} = 2xz\mathbf{i} + (x^2 - y)\mathbf{j} + (2z - x^2)\mathbf{k}.$$
Is it conservative or non-conservative?

17. Prove that a necessary and sufficient condition that $\oint \mathbf{A}\cdot d\mathbf{r} = 0$ for every closed curve C is that $\nabla \times \mathbf{A} = 0$ identically.

PROBLEMS

1. If $\mathbf{a}$ is a constant vector, show that grad $(\mathbf{a}\cdot\mathbf{r}) = \mathbf{a}$.
2. If $\mathbf{a}$ is a constant vector, show that curl $(\mathbf{a}\times\mathbf{r}) = 2\mathbf{a}$.
3. If $\mathbf{a}$ is a constant vector, show that
$$\begin{aligned}(\mathbf{a}\times\nabla)\cdot(\mathbf{b}\times\mathbf{c}) &= (\mathbf{a}\cdot\mathbf{b})(\nabla\cdot\mathbf{c}) + (\mathbf{c}\cdot\nabla)(\mathbf{a}\cdot\mathbf{b})\\ &= (\mathbf{a}\cdot\mathbf{c})(\nabla\cdot\mathbf{b}) - (\mathbf{b}\cdot\nabla)(\mathbf{a}\cdot\mathbf{c})\end{aligned}$$
4. What is the difference between
   (i) $\nabla\Phi$ and $\Phi\nabla$
   (ii) $\nabla\cdot\mathbf{A}$ and $\mathbf{A}\cdot\nabla$
   (iii) $\nabla\times\mathbf{A}$ and $\mathbf{A}\times\nabla$?
5. Find the gradient of $\Phi = x^2y - y^2z - xyz$ at $(1, -1, 0)$ in the direction of vector $\mathbf{i} - \mathbf{j} + 2\mathbf{k}$.
6. Find the unit vector normal to $xyz = 2$ at $(1, -1, -2)$.
7. Find the angle between the surfaces defined by $r^2 = 9$ and $x + y + z^2 = 1$ at point $(2, -2, 1)$.

8. The components of vector **A** are given by

$$A_k = x_i \frac{\partial f}{\partial x_j} - x_j \frac{\partial f}{\partial x_i}, \ (i, j, k) \text{ cyclic and where } f = f(x_1, x_2, x_3)$$

Show that $\mathbf{A} = \mathbf{r} \times \nabla f$, $\mathbf{A}\cdot\mathbf{r} = 0$ and $\mathbf{A}\cdot\nabla f = 0$.

where $\mathbf{r} = \hat{\mathbf{e}}_1 x_1 + \hat{\mathbf{e}}_2 x_2 + \hat{\mathbf{e}}_3 x_3$.

9. Prove the following relations:
   (i) grad $r^n = nr^{n-2}\mathbf{r}$.
   (ii) div $(r^n\mathbf{r}) = (n+3)r^n$.
   (iii) curl $\mathbf{r} = 0$.
   (iv) $\nabla^2 r^n = n(n+1)r^{n-2}$.
   (v) grad $(\ln \mathbf{r}) = \mathbf{r}/r^2$.
   (vi) $\nabla^2(\ln \mathbf{r}) = 1/r^2$.
   (vii) div ($\mathbf{r}$ grad $1/r^3$) $= 3/r^4$.
   (viii) $\nabla^2[\nabla(\mathbf{r}/r^2)] = 2/r^4$.
   (ix) grad $f(\mathbf{r}) = (\mathbf{r}/r)(\partial f/\partial r)$.
   (x) curl $[\mathbf{r}f(\mathbf{r})] = 0$.
   (xi) $\mathbf{A} \times (\nabla \times \mathbf{B}) - (\mathbf{A} \times \nabla) \times \mathbf{B} = \mathbf{A}(\nabla\cdot\mathbf{B}) - (\mathbf{A}\cdot\nabla)\mathbf{B}$.
   (xii) $(\mathbf{A}\cdot\text{grad})A = \frac{1}{2}\,\text{grad}\, A^2 - \mathbf{A} \times \text{curl}\,\mathbf{A}$.
   (xiii) div $(\Phi_1 \text{ grad } \Phi_2 - \Phi_2 \text{ grad } \Phi_1) = \Phi_1\nabla^2\Phi_2 - \Phi_2\nabla^2\Phi_1$.
   (xiv) $\mathbf{A}\,\text{div}\,\mathbf{A} = \frac{1}{2}\,\text{grad}\, A^2 - (\mathbf{A} \times \nabla) \times \mathbf{A}$.
10. Show that the divergence operator is invariant under a rotation of coordinates.
11. A particle rotates with angular velocity ω about an axis which passes through origin. Position vector **r** of the particle has velocity $\dot{\mathbf{r}} = \mathbf{V}$. Calculate curl **V** and div **V** when (a) ω = constant, and (b) $\omega = k/r^2$, where k is constant.
12. A vector field is defined by $\mathbf{V} = f(r)\mathbf{r}$. Show that
    (a) $f(r) = \text{constant} \times r^{-3}$, if $\nabla\cdot\mathbf{V} = 0$, and
    (b) $\nabla \times \mathbf{V} = 0$.
13. For the spherical charge distributions given by
    (i) $\rho = \rho_0 e^{-r/a}$, and
    (ii) $\rho = \rho_0\left(1 - \dfrac{r^2}{a^2}\right)$, for $r \leqslant a$
    $= 0$, for $r > a$,

    find the electric field intensity and potential at any r.
    [In case (ii) find these quantities for $r > a$ only.]
14. If **A** and **B** are irrotational, show that $\mathbf{A} \times \mathbf{B}$ is solenoidal
15. If **A** is irrotational, show that $\mathbf{A} \times \mathbf{r}$ is solenoidal.
16. Vector function $\mathbf{F}(x, y, z)$ is not irrotational, but the product of **F** and scalar function $g(x, y, z)$ is irrotational. Comment.
17. Show that $\mathbf{F}\cdot\nabla \times \mathbf{F} = 0$.
18. Prove that $(\nabla u) \times (\nabla v)$ is solenoidal if u and v are differentiable scalar functions.

19. Scalar function Φ satisfies Laplace's equation $\nabla^2\Phi = 0$. Show that $\nabla\Phi$ is both solenoidal and irrotational.
20. Velocity of a two-dimensional flow of a liquid is given by $\mathbf{V} = \mathbf{i}u(x, y) - \mathbf{j}v(x, y)$. If the liquid is incompressible and the flow is irrotational, show that
$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$
[These are the Cauchy-Riemann conditions.]
21. Prove that
(a) $\nabla \times (\psi\,\nabla\psi) = 0$, and
(b) $(\mathbf{r} \times \nabla)(\mathbf{r} \times \nabla)\psi = r^2\nabla^2\psi - r^2\dfrac{\partial^2\psi}{\partial r^2} - 2r\dfrac{\partial\psi}{\partial r}$,
where ψ is a scalar function.
22. Show that any solution of the equation
$$\nabla \times \nabla \times \mathbf{A} - k^2\mathbf{A} = 0$$
automatically satisfies the vector Helmholtz equation $\nabla^2\mathbf{A} + k^2\mathbf{A} = 0$ and the solenoidal condition $\nabla\cdot\mathbf{A} = 0$.
23. Compute the integral $\int_c \mathbf{V}\cdot d\mathbf{r}$, where $\mathbf{V} = \mathbf{i}x - \mathbf{j}y + \mathbf{k}z$ over the helical path
$$x = \cos t, \quad y = \sin t \quad \text{and} \quad z = t,$$
joining the points determined by $t = 0$ and $t = \frac{\pi}{2}$ and also when c is a straight line joining these points.
24. Prove that $\int_\sigma \mathbf{r}\cdot d\boldsymbol{\sigma} = 3\tau$, where $\mathbf{r}$ is the position vector of a point on surface σ which encloses volume τ.
25. Find $\int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma}$, where $\mathbf{V} = r^2\mathbf{r}$ and σ is the surface of the sphere $x^2 + y^2 + z^2 = a^2$. Compute the integral directly and also by using the divergence theorem.
26. If $\mathbf{V} = \text{grad}\,\Phi$ and $\nabla^2\Phi = \rho$, a specified scalar point function, show that
$$\int_\sigma \frac{\partial\Phi}{\partial n}\,d\sigma = \int_\tau \rho\,d\tau.$$
$\left[\text{Hint: Use } \dfrac{\partial\Phi}{\partial n} = \nabla\Phi\cdot\hat{\mathbf{n}}\right]$.
27. Evaluate the integral $\int \mathbf{F}\cdot d\mathbf{r}$, if $\mathbf{F} = -\mathbf{i}x + \mathbf{j}y/x^2 + y^2$ between $(-1, 0)$ and $(1, 0)$ for
(a) a circular path along a semicircle with origin as centre, and
(b) along the straight lines $(-1, 0)$, $(0, 1)$, $(1, 0)$.

3

Mechanics of a Single Particle and of Systems of Particles

Mechanics or the science of motion of material bodies has been a matter of great interest for ages. This science of motion of material bodies involves two aspects—kinematics and dynamics. In kinematics, we describe the motion of the material bodies in terms of quantities such as displacement, velocity, etc. It is a geometrical description of the motion. In dynamics, we investigate the causes of motion and the properties of the moving systems by studying the equation of motion containing force. The name 'classical mechanics' is used in this century to distinguish it from a modern branch of physics called 'quantum mechanics'.

In this chapter, we begin our discussion with the dynamics of a particle. By a particle we mean a point body having some mass or a body of some mass but having negligible physical dimensions in comparison with the dimensions of other bodies involved in the systems under consideration. The concept of a point particle is a mathematical idealisation of an object whose physical dimensions are negligible in a particular description of its motion under consideration. Thus, in the study of the motion of a planet around the sun, the planet can be treated as a particle. However, in some cases, an atom which can go too close to another charged body, such as a molecule, may not be treated as a particle.

The study of particle dynamics can be best developed on the basis of Newton's laws of motion. These laws are ingenious generalizations based on observations and hence are said to be 'empirical laws'. These laws can be looked upon as postulates of classical mechanics.

3.1 NEWTON'S LAWS OF MOTION

Newton's laws of motion are stated in the following form:

(i) "*Every body continues to be in its state of rest or of uniform motion*

in a straight line unless it is compelled to change that state by external forces acting on it."

(ii) "*The time rate of change of momentum of a particle is proportional to the external force and is in the direction of the force.*"

(iii) "*To every action there is always an equal and opposite reaction*" or "*the mutual actions of any two bodies are always equal and oppositely directed along the same straight line.*"

We shall try to understand the meaning of the laws by using our experimental knowledge as a guide. The first law contains the words '*body*' and '*straight line*'. We shall take the body to be equivalent to a particle and shall accept the usual geometrical concept of a straight line. The first law introduces two important concepts, namely (a) state of rest or of uniform motion, and (b) force. The concept of state of rest or of uniform motion are kinematical concepts. A body may appear to be at rest to one observer, but it may be in motion for some other observer. Since Newton's laws are based on physical observations, the concept of a 'frame of reference' from which the observations are made is automatically associated with the laws.

The property of a body to remain in the state of rest or of uniform motion when no external force acts on it is known as inertia and a frame of reference in which such a state is observed is known as the inertial frame of reference. We shall consider later the frames of reference in greater details.

The other concept introduced is of a force. Newton's first law of motion tells us about the motion of a body when no force acts on it. This law does not tell us what the force does; but it simply tells us what happens when it is absent. One can interpret the first law as the definition of '*zero force*'.

The meaning of the force in terms of the changes it produces in the momentum

$$\mathbf{p} = m\mathbf{v} \tag{3.1}$$

is given by Newton's second law which can be expressed as

$$\mathbf{F} \propto \frac{d\mathbf{p}}{dt}$$

or

$$\mathbf{F} = k\frac{d\mathbf{p}}{dt} = k\frac{d}{dt}(m\mathbf{v}) = km\frac{d\mathbf{v}}{dt}$$

where k is the constant of proportionality. This constant can be chosen to be equal to unity by defining the unit of the force as that force which while acting on a body of unit mass produces a unit acceleration. We have introduced a quantity mass and assumed it to be a constant, which may not always be true. Thus, the expression of Newton's second law becomes

$$\mathbf{F} = \frac{d}{dt}(m\mathbf{v}) = m\frac{d\mathbf{v}}{dt} \tag{3.2}$$

From the second law, it follows that when external force $\mathbf{F} = 0$, momentum $m\mathbf{v}$ is a constant, i.e., the body will continue to be in the state of uniform motion. We can consider the state of rest as a special case of state of uniform motion when $\mathbf{v} = 0$. Thus, the first law is a special case of the second law.

One is surprised at the lack of content in the first law and wonders why Sir Isaac Newton has given it the status of a law. The superfluous nature of the first law has been pointed out by Sir Arthur Eddington by saying that 'every particle continues in its state of rest or of uniform motion except in so far as it does not'. Such a comment, however, would be unfair to Newton who might have some definite idea in the statement of the first law.

We are familiar with a simple experiment in which a ball is allowed to roll along a smooth horizontal surface. By making the surface smoother and smoother, we believe that the ball will continue to move even upto an infinitely long distance with constant momentum if the friction is completely eliminated. If, however, one looks at the motion of heavenly bodies which are moving through space without any resistance to their motion, one finds that they move along curved paths. In fact, in our simple experiment the horizontal surface we imagine is also a part of the spherical surface of the earth. Thus, a large scale uniform motion along a straight line (here we neglect cases like dropping stone, rain drop, etc.) is a rarer phenomenon than motion along a curved path. This is why Greek philosophers considered motion along a circle as 'perfect motion'. That motion of a particle under zero force will be along a straight line is no doubt a great generalization from simple observations of small-scale motion. It was, possibly, the necessity to emphasize this concept, which is now of historical importance, that might have led Newton to give the statement the status of a law.

In the discussion on force we were considering the external force acting on a body. If more forces $\mathbf{F}_1, \mathbf{F}_2, \ldots$ are acting on the particle, then the total effect in motion produced by these forces can be looked upon as produced by a single force $\mathbf{F}$ which is the vector sum of all these forces. This principle is known as the 'principle of superposition' and is one of the fundamental principles in physical theories.

The third law applies only to the two isolated particles exerting forces on each other when forces due to all other particles are completely absent. The forces of action and reaction or of mutual interaction between the two particles do so along the line joining the two particles. According to the third law, the force acting on particle 1 due to particle 2, viz. $\mathbf{F}_1$ is equal and opposite to the force acting on particle 2 due to particle 1, viz. $\mathbf{F}_2$. Thus, $\mathbf{F}_1 = -\mathbf{F}_2$.

Equation (3.2), viz. $\mathbf{F} = m\mathbf{a}$ can be used to measure the mass of a body. Thus, let the same force $\mathbf{F}$ act on two bodies having masses m_1 and m_2 so

as to produce accelerations $\mathbf{a}_1$ and $\mathbf{a}_2$ respectively. Thus, we can write

$$m_1\mathbf{a}_1 = m_2\mathbf{a}_2$$

or $$m_1 a_1 = m_2 a_2 \quad \text{numerically.}$$

Hence, $$\frac{m_1}{m_2} = \frac{a_2}{a_1} \tag{3.3}$$

If, now, m_2 is known or taken as unit mass, m_1 can be calculated by measuring $\frac{a_2}{a_1}$. The mass of a body measured in this manner is called the inertial mass.

Another way of measuring the mass of a body is by weighing it, i.e., by comparing the gravitational force acting on it with that on a standard mass. In this procedure, we make use of the fact that in the gravitational field the weight of a body is exactly equal to the gravitational force acting on it. In that case, equation $\mathbf{F} = m\mathbf{a}$ becomes the weight, $\mathbf{W} = m\mathbf{g}$, where g is the acceleration due to gravity. The mass of a body measured in this manner is called the gravitational mass.

In recent experiments performed to find whether these two masses are identical or not it has been established that inertial and gravitational masses are equal within an error of a few parts in 10^{10}. The assertion of the exact equality of the inertial and the gravitational mass is called the *weak principle of equivalence.*

Frames of Reference

The concept of absolute rest or of uniform motion introduced in Newton's first law involves the concept of a frame of reference with respect to which the state of rest or of uniform motion is observed. Consider two frames of reference S and S' moving with relative velocity $\mathbf{v}$, say along x- (or x'-) axis for simplicity (Fig. 3.1). The observers in the two frames will give two different coordinates to the same particle at point

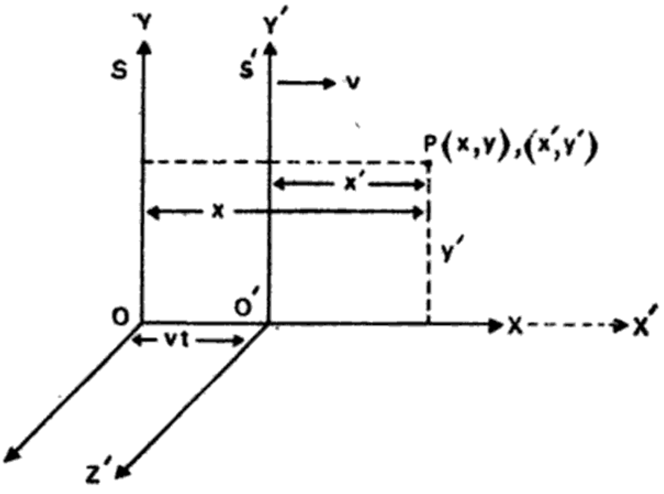


Fig. 3.1 Frames of reference S and S' in uniform relative motion along the x-axis

P which is observed by both. The coordinates are related to each other by transformation equations known as the *Galilean transformations*. Thus,

$$\begin{aligned} x' &= x - vt \\ y' &= y \\ z' &= z \\ \text{and} \qquad t' &= t \end{aligned} \tag{3.4}$$

Here, we are assuming that initially (i.e. at $t = 0$), the two coordinate systems were coincident and the motion of S' is along the direction of the x-axis only with uniform velocity $\mathbf{v}$. It is clear that Newton's second law of motion will have the same form in these two frames since $d^2x'/dt'^2 = d^2x/dt^2$. Hence, the second law of motion is said to be invariant with respect to the Galilean transformation. Such non-accelerated frames of reference—non-accelerated with respect to one another and in which law of inertia holds—are called *inertial frames of reference*. If the transformation equations contain an acceleration term, i.e., if frame S' is accelerated with respect to S, then the form of the law of motion would be different in the two frames and the nature of the phenomenon observed by the two observers would be different.

Is it possible to have an inertial frame of reference? A frame of reference fixed to the earth is obviously not an inertial frame since the earth is rotating about its own axis. We can then fix the frame at the centre of the earth. But, the earth is revolving around the sun. One may then fix the frame at the centre of the sun. But, the sun is also not at rest since it is accelerated in the galaxy. In this way, the search for a *point at rest* can be extended. Before the advent of Einstein's Theory of Relativity, it was considered possible to find out a star in the universe which might be at *absolute rest*. A frame of reference fixed on such a *fixed star* may be considered an absolute frame where we can observe *absolute rest* and any motion measured with respect to this frame can be considered an *absolute motion*.

Newton's laws, however, do not refer to *absolute rest* or *absolute uniform motion* and hence do not imply the necessity of a '*fixed star*' frame. The laws demand that if no force acts, the body should move with constant velocity in a reference frame called an inertial frame. Amongst the two frames moving relatively to each other with constant velocity, no particular frame is more important and any one can be used to solve dynamical problems. It is very difficult to have a perfectly inertial frame since bodies like the earth, the moon, the sun, etc. are moving along curved paths and hence are accelerated. We have, therefore, to select a frame of reference for a dynamical problem which is very close to an inertial frame, i.e., acceleration of the frame should be negligibly small as compared to the acceleration involved in the motion under study.

In considering an inertial frame, we implicitly require an Euclidian space, i.e., the usual three-dimensional space, and the assumption that this

space is homogeneous and isotropic. By this we mean that the phenomena observed at various points will be identical and will not depend on the orientation of the coordinate axes. We also use the concept of homogeneity or uniformity of time, i.e., a particle moving with a constant velocity or a particle on which no external force is acting will traverse equal intervals of space in equal intervals of time. In the Galilean transformations, time is independent of the frame of reference chosen, i.e. $t' = t$ whatever may be the motion of the frame of reference. This, as will be seen in Chapter 14, is a consequence of the fact that a signal, say a flash of light sent out to observe the position of a body, is assumed to have infinitely large velocity. This situation changes in the Special Theory of Relativity in which time is different in different frames of reference moving relatively to each other, since the velocity of light is assumed to be finite and constant, a fact well known experimentally.

3.2 MECHANICS OF A PARTICLE

The essential problem in mechanics, developed with Newton's laws of motion as the basis, is to solve the differential equation given by Newton's second law of motion viz.

$$\mathbf{F} = m\mathbf{a}$$

or

$$\mathbf{F} = m\frac{d^2\mathbf{r}}{dt^2} \tag{3.5}$$

under given initial values of position and velocity of the particle on which force $\mathbf{F} = \mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}, t)$ is acting. The solution of equation (3.5) is of the form

$$\mathbf{r} = \mathbf{r}(t) \tag{3.6}$$

Equation (3.6) describes a certain path and gives the dependence of the position vector $\mathbf{r}$ on time t. It is not always possible to find explicit solutions of equations (3.5) and we shall only restrict ourselves to simpler problems wherein exact or at least approximate solution can be obtained.

If no external force is acting on the particle, then

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = 0 \tag{3.7a}$$

or

$$\mathbf{p} = m\mathbf{v} = \text{const} \tag{3.7b}$$

Thus, if there is no force acting on the particle, the momentum of the particle is constant. If the mass of the particle is assumed to remain constant, it will continue to move along a straight line. This corresponds to a statement of Newton's first law of motion.

Angular momentum $\mathbf{L}$ of the particle about point O is defined as

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \tag{3.8}$$

where $\mathbf{r}$ is the position vector and $\mathbf{p}$ is the linear momentum of the particle at the given instant.

The time rate of change of angular momentum **L** is defined as torque **N**. Thus,

$$\mathbf{N} = \frac{d\mathbf{L}}{dt} = \frac{d}{dt}(\mathbf{r} \times \mathbf{p}) = \dot{\mathbf{r}} \times \mathbf{p} + \mathbf{r} \times \dot{\mathbf{p}}$$

But,
$$\dot{\mathbf{r}} \times \mathbf{p} = \mathbf{v} \times m\mathbf{v} = 0$$

and
$$\mathbf{r} \times \dot{\mathbf{p}} = \mathbf{r} \times \frac{d\mathbf{p}}{dt} = \mathbf{r} \times \mathbf{F}$$

Hence,
$$\mathbf{N} = \frac{d\mathbf{L}}{dt} = \mathbf{r} \times \mathbf{F} \tag{3.9}$$

If torque **N** acting on the particle is zero, we have $\frac{d\mathbf{L}}{dt} = 0$ or $\mathbf{L} =$ a constant. Thus, if no torque is acting on a particle, its angular momentum is constant. Planets moving around the sun is the finest example of this conservation law.

Work Done The work done by the total external force in moving a particle from position 1 to position 2 is given by

$$W_{12} = \int_1^2 \mathbf{F} \cdot d\mathbf{r} \tag{3.10}$$

If the mass of the particle is constant, we can write

$$\begin{aligned} W_{12} &= \int_1^2 m \frac{d\mathbf{v}}{dt} \cdot d\mathbf{r} \\ &= m \int_1^2 \frac{d\mathbf{v}}{dt} \cdot \frac{d\mathbf{r}}{dt}\, dt \\ &= m \int_1^2 \left(\frac{d\mathbf{v}}{dt} \cdot \mathbf{v}\right) dt \\ &= \tfrac{1}{2} m \int_1^2 d(v^2) \\ &= \tfrac{1}{2} m (v_2^2 - v_1^2) \\ &= T_2 - T_1 \end{aligned} \tag{3.11}$$

where T_1 and T_2 are the kinetic energies of the particle in positions 1 and 2 respectively.

If $T_1 > T_2$, $W_{12} < 0$, i.e., work is done by the particle against the force and its kinetic energy has decreased.

If $T_2 > T_1$, $W_{12} > 0$, i.e., work is done by the force on the particle and the kinetic energy of the particle has increased.

In any case, the work done depends upon the difference in kinetic energies of the particle in the two positions. The work done against dissipative forces like the frictional force is always negative.

If the force-field is such that the work done along a closed path is zero, then the force is said to be *conservative*.

Thus, if $\oint \mathbf{F} \cdot d\mathbf{r} = 0$, then the force **F** is a conservative force.

We now convert this line integral into a surface integral by using

Stokes's theorem. Then, we get

$$\oint \mathbf{F}\cdot d\mathbf{r} = \int_\sigma \nabla \times \mathbf{F}\cdot d\boldsymbol{\sigma} \tag{3.12}$$

Thus, for a conservative force-field

$$\int_\sigma \nabla \times \mathbf{F}\cdot d\boldsymbol{\sigma} = 0 \tag{3.13}$$

Since $d\boldsymbol{\sigma}$, the surface element is arbitrary, we must have

$$\nabla \times \mathbf{F} = 0 \tag{3.14}$$

Equation (3.14) is the necessary and sufficient condition for force $\mathbf{F}$ to be conservative. We know that the curl of the gradient of a scalar point function is zero. Hence, we can write

$$\mathbf{F} = -\text{grad } V \tag{3.15}$$

Scalar point function V introduced in equation (3.15) is called the potential energy of the particle at that point. The negative sign on the right-hand side indicates that $\mathbf{F}$ is in the direction of decreasing V. We can write the expression for W_{12} as

$$\begin{aligned} W_{12} &= \int_1^2 \mathbf{F}\cdot d\mathbf{r} = -\int_1^2 \nabla V\cdot d\mathbf{r} \\ &= -\int_1^2 \left(\frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz\right) \\ &= -\int_1^2 dV \\ &= V_1 - V_2 \end{aligned} \tag{3.16}$$

Thus, the work done by the force in displacing the particle from position 1 to position 2 is equal to the difference between the potential energies of the particle in those two positions.

Equating equations (3.11) and (3.15), we get

$$T_1 + V_1 = T_2 + V_2 = \text{const} = E \tag{3.17}$$

Thus, the sum of potential and kinetic energies of a particle at every point in a conservative force-field is constant. Gravitational and electrostatic fields are the common examples of conservative fields.

Potential energy V introduced through equation (3.15) makes it clear that it is not unique, and an addition of any constant energy C to it does not change the equation $\mathbf{F} = -\text{grad } V$. Thus,

$$\mathbf{F} = -\text{grad } V = -\text{grad } (V + C) \tag{3.18}$$

since C is constant. Hence, the absolute value of the potential energy has no meaning. We can determine only the differences in potential energies as in equation (3.16).

It should be remembered that the introduction of potential energy is a convenient device to describe a force-field by a scalar function. Equation of motion (3.2) which describes the physical situation contains $\mathbf{F}$ and not V. The equation of motion can also be written through V as will be done

later in Lagrange's and Hamilton's equations of motion.

The determination of the kinetic energy is also relative. As mentioned earlier, we use a certain inertial frame of reference for measuring the velocity and hence the kinetic energy. Consider two inertial frames of reference S and S' moving with relative constant velocity $\dot{\mathbf{R}}$ with respect to each other (Fig. 3.2).

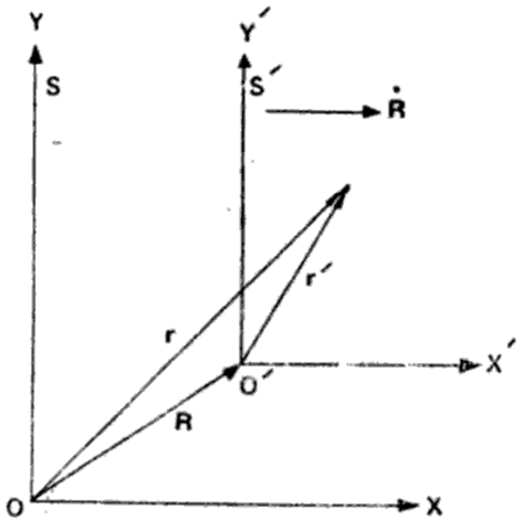


Fig. 3.2 Frames of reference in uniform relative motion (general case)

Let $\mathbf{r}$ and $\mathbf{r}'$ be the position vectors of particle P with respect to S and S' respectively. Then, $\dot{\mathbf{r}}$ and $\dot{\mathbf{r}}'$ are the corresponding velocities of the particle. These values will be related to each other by the equations

$$\mathbf{r} = \mathbf{R} + \mathbf{r}' \tag{3.19}$$

and

$$\dot{\mathbf{r}} = \dot{\mathbf{R}} + \dot{\mathbf{r}}' \tag{3.20}$$

Substitution of $\mathbf{r}$ or $\dot{\mathbf{r}}$ from equations (3.19) and (3.20) shows that the equations of motion, viz. (3.5) or (3.9), have the same form.

As we are using an inertial frame of reference for measuring velocity, the kinetic energy is relative. The absolute kinetic energy could have been measured if we could find a frame of reference which is absolutely at rest. It is, however, impossible to find such a frame of reference in accordance with the Special Theory of Relativity. Hence, the concept of absolute kinetic energy is meaningless. The total energy of a particle (or of a system of particles), viz. $E = T + V$, is not known in the absolute sense. Such a knowledge is unnecessary too and only the differences in the energies are of physical significance.

Consider the time derivative of total energy E,

i.e.
$$\frac{dE}{dt} = \frac{dT}{dt} + \frac{dV}{dt}$$

But,
$$\mathbf{F} \cdot d\mathbf{r} = m \frac{d\mathbf{v}}{dt} \cdot \mathbf{v}\, dt = \frac{d}{dt} (\tfrac{1}{2} m v^2)\, dt = dT$$

or
$$\mathbf{F} \cdot \frac{d\mathbf{r}}{dt} = \frac{dT}{dt}$$

Further,
$$dV = \frac{\partial V}{\partial x}\,dx + \frac{\partial V}{\partial y}\,dy + \frac{\partial V}{\partial z}\,dz + \frac{\partial V}{\partial t}\,dt$$
$$= (\nabla V \cdot d\mathbf{r}) + \frac{\partial V}{\partial t}\,dt$$

Hence,
$$\frac{dV}{dt} = \left(\nabla V \cdot \frac{d\mathbf{r}}{dt}\right) + \frac{\partial V}{\partial t}\,dt$$

Substituting these values, we get
$$\frac{dE}{dt} = \mathbf{F} \cdot \frac{d\mathbf{r}}{dt} + \left(\nabla V \cdot \frac{d\mathbf{r}}{dt} + \frac{\partial V}{\partial t}\right)$$

or
$$\frac{dE}{dt} = [\mathbf{F} + \nabla V] \cdot \frac{d\mathbf{r}}{dt} + \frac{\partial V}{\partial t} \tag{3.21}$$

The first term on the right-hand side of equation (3.21) will vanish if $\mathbf{F} = -\nabla V$, and
$$\frac{dE}{dt} = \frac{\partial V}{\partial t} \tag{3.22}$$

If, furthermore, the potential energy is not an explicit function of time, $\frac{\partial V}{\partial t} = 0$, then
$$\frac{dE}{dt} = 0 \quad \text{or} \quad E = \text{const}$$

Thus, the total energy of the particle moving in a conservative force-field remains constant, if the potential is not an explicit function of time.

3.3 EQUATION OF MOTION OF A PARTICLE

In general, the force acting on a particle may depend on its position, velocity and time and hence the equation of motion of the particle is
$$m\ddot{\mathbf{r}} = \mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}, t) \tag{3.23}$$

This is a second order differential equation in space coordinates and after integrating it twice, we will get the trajectory of the particle. There will be two constants of integration introduced and these will be determined by knowing the initial conditions of the particle. Integrating equation (3.23) with respect to time, we get
$$m\int_{t_0}^{t} \ddot{\mathbf{r}}\,dt' = \int_{t_0}^{t} \mathbf{F}\,dt'$$

or
$$m(\dot{\mathbf{r}} - \mathbf{v}_0) = \int_{t_0}^{t} \mathbf{F}\,dt' \tag{3.24}$$

where we have taken initially, i.e. at $t = t_0$, velocity $\dot{\mathbf{r}}$ to be equal to $\mathbf{v}_0$. The quantity on the right-hand side of equation (3.24) is called the impulse of the force and the integral represents the impulse imparted to the particle during time interval $(t - t_0)$ and is equal to the change in the momentum of the particle. Integrating equation (3.24) once again with

respect to time, we get

$$\mathbf{r} - \mathbf{r}_0 = \mathbf{v}_0(t - t_0) + \frac{1}{m}\int_{t_0}^{t} dt' \int_{t_0}^{t} \mathbf{F}\, dt'' \tag{3.25}$$

where $\mathbf{r}$ and $\mathbf{r}_0$ are the position vectors of the particle at instants $(t = t)$ and $(t = t_0)$ respectively. If force $\mathbf{F}$ is known and integration of equation (3.25) is carried out, we get the explicit form of equation of the trajectory of the particle. In simple forms of $\mathbf{F}$, this integration is possible. In complicated cases, however, we have to resort to numerical integration. While solving problems in mechanics, it is essential to decide first the body whose motion is to be studied and then to consider all the forces applied to that body alone.

We shall now consider some simple forms of forces that we come across in nature. These are,

(a) $\mathbf{F} = m\mathbf{a}$, a constant. For example $\mathbf{F} = m\mathbf{g}$, when $\mathbf{g}$, the acceleration due to earth's gravity can be considered constant, i.e., near the surface of the earth.

(b) $\mathbf{F} = \mathbf{F}(t)$. For example, alternating force of the type $\mathbf{F} = \mathbf{F}_0 \sin \omega t$, which is experienced by charged particles in alternating fields.

(c) $\mathbf{F} = \mathbf{F}(\mathbf{r})$. For example, $\mathbf{F} = k\mathbf{r}/r^3$, the force given by inverse square law experienced in the gravitational or electrostatic field, or $\mathbf{F} = -k\mathbf{r}$ which is experienced by a particle performing linear simple harmonic motion.

(d) $\mathbf{F} = \mathbf{F}(\dot{\mathbf{r}})$. For example, $\mathbf{F} = k\mathbf{v}$ or $\mathbf{F} = k\mathbf{v}^n$ which represents the frictional force experienced by a particle moving in a viscous medium.

One or more of the above types of forces might be acting on a particle. We shall consider motion under inverse square law and oscillatory motion in later chapters. In this chapter, we shall consider some simple cases of the remaining types of forces.

(a) Motion Under Constant Force

When constant force $\mathbf{F} = m\mathbf{a}$ is acting on a particle, equation (3.24) becomes

$$\dot{\mathbf{r}} = \mathbf{v}_0 + \mathbf{a}(t - t_0) \tag{3.26}$$

Integration of this equation gives, or from equation (3.25), we get

$$\mathbf{r} - \mathbf{r}_0 = \mathbf{v}_0(t - t_0) + \tfrac{1}{2}\mathbf{a}(t - t_0)^2 \tag{3.27}$$

This is the vector equation corresponding to one-dimensional form $s = ut + \frac{1}{2}at^2$ which is quite familiar to the reader.

Atwood's Machine Consider a system of two masses m_1 and m_2 $(m_2 > m_1)$ tied by a light inextensible string of length l. The masses are hanging over a pulley as shown in Fig. 3.3a.

Since the masses can move only in the vertical direction, this is a one-dimensional problem. Let the position of, say m_2 be given by x. This

also fixes the position of m_1. As $m_2 > m_1$, velocity $\frac{dx}{dt}$ of m_2 will be downwards whereas that of m_1 will be upwards. We neglect the friction between the string and the pulley.

The force acting on m_1 moving upwards is

$$\mathbf{F}_1 = \boldsymbol{\tau} - m_1\mathbf{g} \tag{3.28}$$

and that on m_2 moving downwards is

$$\mathbf{F}_2 = -\boldsymbol{\tau} + m_2\mathbf{g} \tag{3.29}$$

where $\boldsymbol{\tau}$ is the tension in the string. According to Newton's second law

$$m_1\ddot{\mathbf{x}} = \boldsymbol{\tau} - m_1\mathbf{g} \tag{3.30}$$

and

$$m_2\ddot{\mathbf{x}} = -\boldsymbol{\tau} + m_2\mathbf{g} \tag{3.31}$$

Addition of equations (3.30) and (3.31) gives acceleration

$$\ddot{\mathbf{x}} = \frac{m_2 - m_1}{m_1 + m_2}\mathbf{g} \tag{3.32}$$

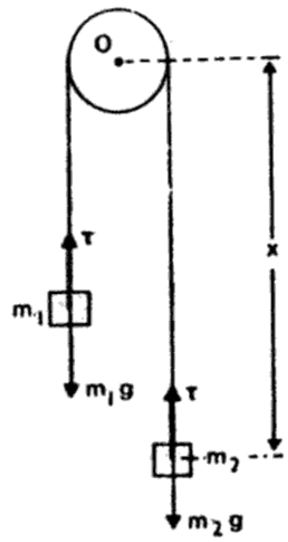


Fig. 3.3a Atwood's machine

Solving equations (3.30) and (3.31) for $\boldsymbol{\tau}$, we get

$$\boldsymbol{\tau} = \frac{2m_1m_2}{m_1 + m_2}\mathbf{g} \tag{3.33}$$

It should be noted that if $m_1 = m_2$, acceleration $\ddot{\mathbf{x}} = 0$ and the two masses are stationary. Similarly, if one of the masses is very large, i.e. $m_2 \gg m_1$, then $\ddot{\mathbf{x}} \simeq \mathbf{g}$ and $\boldsymbol{\tau} \simeq 2m_1\mathbf{g}$ which is negligible as compared to $m_2\mathbf{g}$, and hence mass m_2 undergoes free fall.

(b) Motion Under a Force which Depends on Time Only

Consider time dependent force

$$\mathbf{F} = \mathbf{F}_0 \sin \omega t \tag{3.34}$$

acting on a particle. For simplicity consider only one-dimensional motion along the x-axis. Then, the equation of motion is

$$\ddot{\mathbf{x}} = \frac{\mathbf{F}_0}{m}\sin \omega t \tag{3.35}$$

On integration of equation (3.35) for a particle which has at the time $t = 0$, $\mathbf{x} = \mathbf{x}_0$ and $\mathbf{v} = \mathbf{v}_0$, we get

$$\dot{\mathbf{x}} = \mathbf{v}_0 + \frac{\mathbf{F}_0}{m\omega}(1 - \cos \omega t)$$

$$= \mathbf{v}_0 + \frac{\mathbf{F}_0}{m\omega} - \frac{\mathbf{F}_0}{m\omega}\cos \omega t \tag{3.36}$$

Integrating equation (3.36) once again, we get

$$\mathbf{x} = \mathbf{x}_0 + \left(\mathbf{v}_0 + \frac{\mathbf{F}_0}{m\omega}\right)t - \frac{\mathbf{F}_0}{m\omega^2}\sin \omega t \tag{3.37}$$

This problem is of interest in connection with scattering of electro-

magnetic radiation by free electrons such as those appearing in the ionosphere.

(c) Motion Under a Force Dependent on Distance Only

Motion of a particle moving under the action of gravitational or Coulomb force will be considered in detail in Chapter 5, and that under harmonically varying force in Chapter 6.

(d) Motion of a Particle Subjected to a Resistive Force

1. Let us suppose that a body is projected horizontally with initial velocity $\mathbf{v}_0$ in a medium which offers resistance proportional to the first power of velocity. The equation of motion for the horizontal component of the motion in this case is

$$F = m\frac{dv}{dt} = -kmv \tag{3.38}$$

Let us consider motion along the x-axis only and let k be the resistive force of the medium per unit velocity per unit mass.

From equation (3.38), we get

$$\frac{dv}{v} = -k\,dt \tag{3.39}$$

Integrating equation (3.39), we get

$$\ln v = -kt + C_1 \tag{3.40}$$

If initially $t = 0$, $v = v_0$, then

$$\ln v_0 = C_1$$

Then, equation (3.40) becomes

$$\ln\frac{v}{v_0} = -kt$$

or

$$v = v_0 e^{-kt} \tag{3.41}$$

Integrating equation (3.41) again we will get the trajectory of the body. Thus, equation (3.41) is

$$\frac{dx}{dt} = v_0 e^{-kt}$$

Hence,

$$\int dx = v_0 \int e^{-kt}\,dt$$

or

$$x = -\frac{v_0}{k}e^{-kt} + C_2 \tag{3.42}$$

If we choose initial conditions as $x = 0$ at $t = 0$, then

$$C_2 = \frac{v_0}{k}$$

and equation (3.42) becomes

$$x = \frac{v_0}{k}(1 - e^{-kt}) \tag{3.43}$$

and gives the position of the body at any time t. As time elapses, the

exponential factor decreases, i.e. $e^{-kt} \to 0$ as $t \to \infty$, and the body comes to a halt at distance $\frac{v_0}{k}$.

2. Consider the motion of a particle falling under the action of gravity near the surface of the earth. Let us assume that the frictional force of air is proportional to the velocity of the particle. Since the body will be accelerated vertically downwards, we shall take the x-axis along the path of the particle and treat the motion as one-dimensional.

Let the resistive force be $f = mkv$, where m is the mass of the falling body. The equation of motion is

$$F = m\frac{d^2x}{dt^2} = m\frac{dv}{dt} = mg - mkv \tag{3.44a}$$

or

$$\frac{dv}{dt} = g - kv \tag{3.44b}$$

or

$$\frac{dv}{g - kv} = dt \tag{3.45}$$

Let the initial velocity with which the body is dropped be zero, i.e., at $t = 0$, $v = \frac{dx}{dt} = 0$. Integration of equation (3.45) gives

$$-\frac{1}{k}\ln(g - kv) = t + C \tag{3.46}$$

The constant of integration C is obtained from the initial conditions, viz. at $t = 0$, $v = 0$. Thus,

$$-\frac{1}{k}\ln g = C$$

Putting this constant in equation (3.46), we get

$$-\frac{1}{k}[\ln(g - kv) - \ln g] = -\frac{1}{k}\ln\frac{g - kv}{g} = t$$

Hence,

$$g - kv = ge^{-kt}$$

or

$$v = \frac{dx}{dt} = \frac{g}{k}(1 - e^{-kt}) \tag{3.47}$$

This equation gives the velocity of the body in a viscous medium when frictional force is mkv. If the path is long, after a sufficiently long time (as $t \to \infty$, $e^{-kt} \to 0$), the velocity of the body becomes constant and is g/k. This is called the 'terminal velocity'. From equation (3.44a) it is clear that, when the gravitational force is balanced by force of resistance of the medium, the body ceases to have acceleration and $\frac{dv}{dt} = 0$. After a long time the body has the terminal velocity given by

$$g - kv_t = 0$$

or

$$v_t = \frac{g}{k} \tag{3.48}$$

Integrating equation (3.47), we get

$$x = \frac{g}{k}\left(t + \frac{e^{-kt}}{k}\right) + C'$$

To evaluate constant C', let the body be released at point $x = 0$ at time $t = 0$. Then,

$$C' = -\frac{g}{k^2}$$

Hence, the position of the body is given by

$$x = \frac{gt}{k} - \frac{g}{k^2}(1 - e^{-kt}) \tag{3.49}$$

3. If the resistance of the medium is proportional to the square of the velocity, then the equation of motion of a vertically falling body is

$$m\frac{dv}{dt} = mg - mk^2v^2 \tag{3.50}$$

where for convenience we have put resisting force mk^2v^2. Then

$$\frac{dv}{g - k^2v^2} = dt \tag{3.51}$$

This equation can be written as

$$\frac{1}{2k\sqrt{g}}\left(\frac{1}{\sqrt{g} + kv} + \frac{1}{\sqrt{g} - kv}\right) d(kv) = dt \tag{3.52}$$

Integration gives

$$\frac{1}{2k\sqrt{g}} \ln \frac{\sqrt{g} + kv}{\sqrt{g} - kv} = t + C_1 \tag{3.53}$$

With initial conditions $t = 0$, $v = 0$, we get

$$C_1 = 0$$

Hence,
$$kv = \sqrt{g}\,\frac{e^{2kt\sqrt{g}} - 1}{e^{2kt\sqrt{g}} + 1} = \sqrt{g} - \frac{2\sqrt{g}}{1 + e^{2kt\sqrt{g}}} \tag{3.54}$$

Second integration gives

$$kx = \sqrt{g}t - 2\sqrt{g}t + \frac{1}{k}\ln(1 + e^{2kt\sqrt{g}}) + C_2 \tag{3.55}$$

Let $x = 0$ at $t = 0$, then

$$C_2 = -\frac{\ln 2}{k}$$

Thus,
$$x = \frac{1}{k^2}\ln\left(\frac{1 + e^{2kt\sqrt{g}}}{2}\right) - \frac{\sqrt{g}t}{k} \tag{3.56}$$

which gives the position of the body.

4. *Motion of a projectile—no resistance*: Let a body be projected at angle α with the horizontal with velocity $\mathbf{v}_0$. The motion will remain in the vertical plane of velocity vector $\mathbf{v}_0$. Let us take the x-axis along the horizontal and the y-axis upward along the vertical direction in the plane of motion. We write the initial conditions as

$$\left.\begin{array}{l} x(t=0)=0, \quad y(t=0)=0 \\ \dot{x}(t=0)=v_0 \cos \alpha = U \\ \dot{y}(t=0)=v_0 \sin \alpha = V \end{array}\right\} \tag{3.57}$$

The equations of motion of the projectile are

$$m\ddot{x} = 0, \; m\ddot{y} = -mg \tag{3.58}$$

Integration of equations (3.58) gives

$$\dot{x} = c_1, \; \dot{y} = -gt + c_2$$

With the initial conditions given in equations (3.57), we get $c_1 = U$ and $c_2 = V$. This gives

$$\dot{x} = U \text{ and } \dot{y} = -gt + V \tag{3.59}$$

Integrating again, we get

$$x = Ut, \quad y = Vt - \tfrac{1}{2}gt^2$$

These equations together give the equation of the trajectory of the projectile. It is usually expressed in terms of α, the angle of projection. Thus, eliminating t and remembering $\frac{V}{U} = \tan\alpha$, we get

$$y = x \tan\alpha - \frac{gx^2}{2v_0^2}(1 + \tan^2\alpha) \tag{3.60}$$

Thus, the trajectory of projectile is parabola. The range R of the projectile is obtained from equation (3.60) by putting $x = R$ and $y = 0$. Thus,

$$R \tan\alpha = \frac{gR^2}{2v_0^2}(1 + \tan^2\alpha)$$

or

$$R = \frac{v_0^2 \sin 2\alpha}{g} \tag{3.61}$$

The time required to cover the distance equal to range is

$$T = \frac{2V}{g} \tag{3.62}$$

5. *Motion of a projectile in a resisting medium*: Consider the motion of a projectile in the atmosphere in which the retarding force is offered by the air resistance. Let us assume that the retarding force is proportional to the instantaneous velocity. In this case, the equations of motion of the projectile are, in the component form

$$m\ddot{x} = -km\dot{x} \tag{3.63a}$$

and

$$m\ddot{y} = -mg - km\dot{y} \tag{3.63b}$$

where the symbols have the same meanings as those in the last article. Equations (3.63a) and (3.63b) show that the present problem is equivalent to combinations of the problems discussed in 1 and 2 of this article with different initial conditions, i.e., those of 4.

The solution of equation (3.63a) is

$$x = \frac{U}{k}(1 - e^{-kt}) \tag{3.64a}$$

Similarly, the solution of equation (3.63b) can be written aown by evaluating the constants of integration under the new initial conditions. The solution is

$$y = -\frac{gt}{k} + \frac{kV + g}{k^2}(1 - e^{-kt}) \tag{3.64b}$$

To find range R', we find time T required by the projectile for the entire trajectory. For this we note that $y = 0$ at the end of the trajectory. Hence, from equation (3.64b), we get

$$T = \frac{kV + g}{gk}(1 - e^{-kT}) \tag{3.65a}$$

This gives time T in terms of e^{-kT} which can be expanded in a series and can be simplified as follows:

$$T = \frac{kV + g}{gk}[1 - (1 - kT + \tfrac{1}{2}k^2T^2 - \tfrac{1}{6}k^3T^3 \ldots)]$$

$$= \frac{kV + g}{gk}[kT - \tfrac{1}{2}k^2T^2 + \tfrac{1}{6}k^3T^3 \ldots]$$

Simplifying this equation, we get

$$1 = \frac{kV + g}{g}\left[1 - \tfrac{1}{2}kT + \tfrac{1}{6}k^2T^2 \ldots\right]$$

$$= \frac{kV + g}{g} - \left(\frac{kV + g}{g}\right)\frac{kT}{2} + \left(\frac{kV + g}{g}\right)\frac{k^2T^2}{6} \ldots$$

Hence,

$$T\left(1 + \frac{kV}{g}\right) = \frac{2}{k}\left[\left(\frac{kV + g}{g} - 1\right) + \left(\frac{kV + g}{g}\right)\frac{k^2T^2}{6} \ldots\right]$$

$$= \frac{2V}{g} + \left(1 + \frac{kV}{g}\right)\frac{kT^2}{3} \ldots \tag{3.65b}$$

This equation in the form of a series gives the value of the total time of flight T and cannot be solved exactly. We have, therefore, to obtain its approximate value by making successive approximations under the assumption that the terms containing higher orders of kT contribute less and less. For this assumption to be true, obviously, $kT < 1$.

Equation (3.65b) shows that if $k \to 0$, the time of flight of the projectile is

$$T_0 = \frac{2V}{g} \tag{3.65c}$$

Now, suppose that k is very small (but not equal to zero), then the time of flight will approach value T_0. We shall substitute this value in the second-order term and neglect the contributions from the rest. Thus, we get

$$T\left(1 + \frac{kV}{g}\right) = \frac{2V}{g} + \left(1 + \frac{kV}{g}\right)\frac{k}{3}\left(\frac{2V}{g}\right)^2$$

or

$$T \simeq \frac{2V}{g}\left[\left(1 + \frac{kV}{g}\right)^{-1} + \frac{2kV}{3g}\right]$$

$$\simeq \frac{2V}{g}\left[\left(1 - \frac{kV}{g} + \frac{k^2V^2}{g^2} \ldots\right) + \frac{2kV}{3g}\right]$$

Neglecting the terms in k^2 and of higher orders in k and simplifying the equation, we get

$$T \simeq \frac{2V}{g}\left(1 - \frac{kV}{3g}\right) \tag{3.65d}$$

Let us now simplify equation (3.64a) as

$$x = \frac{U}{k}\,[1 - e^{-kt}]$$

$$= \frac{U}{k}\,[kt - \tfrac{1}{2}k^2t^2 + \tfrac{1}{6}k^3T^3 \ldots]$$

If we now substitute the value of T as given by equation (3.65d), we get $x_{(t=T)} = R'$, the range of the projectile in the presence of the air resistance. Thus, restricting to the first two terms, we have

$$R' = U[T - \tfrac{1}{2}kT^2]$$

$$\simeq \frac{2UV}{g}\left[1 - \frac{4kV}{3g}\right]$$

But, $$\frac{2UV}{g} = \frac{2v_0 \cos\alpha\, v_0 \sin\alpha}{g} = \frac{v_0^2 \sin 2\alpha}{g} = R$$

Hence, $$R' = R\left[1 - \frac{4kV}{3g}\right] \tag{3.66a}$$

Therefore, the change in the range of the projectile correct up to first order of k is given by

$$\Delta R = R' - R = -\frac{4kVR}{3g}$$

or $$\Delta R = -\frac{4kv_0^3 \sin\alpha \sin 2\alpha}{3g} \tag{3.66b}$$

After finding the time of flight and the range of the projectile, let us study the nature of the trajectory. For any position of the projectile such as P (Fig. 3.3b) moving with velocity **v** and making angle φ with the

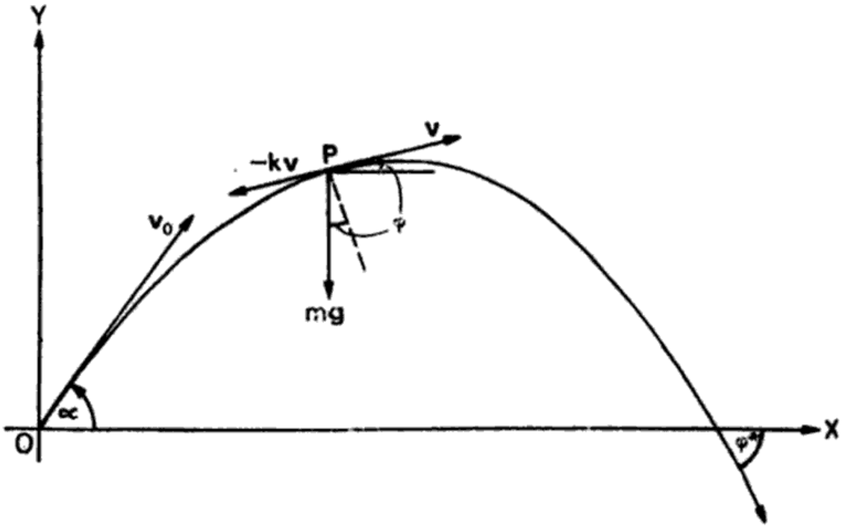


Fig. 3.3b Trajectory of a projectile under the resistive force

horizontal, we can write down the equation of motion by considering the motion along a small arc-length of the trajectory at P. Let the radius of curvature at P be ρ. The equations of motion written in a component form along tangential and radial directions are

$$m\dot{v} = -kv - mg \sin \varphi \tag{3.67}$$

and

$$\frac{mv^2}{\rho} = mg \cos \varphi \tag{3.68}$$

The right-hand sides of equations (3.67) and (3.68) are the components of forces along tangential and radial directions. The radius of curvature is given by

$$\frac{1}{\rho} = -\frac{d\varphi}{ds} = -\frac{d\varphi}{dt}\frac{dt}{ds} = -\frac{1}{v}\frac{d\varphi}{dt}$$

where $ds = \sqrt{dx^2 + dy^2} = v\,dt$.

Equation (3.68), after substituting for ρ becomes

$$v\frac{d\varphi}{dt} = -g \cos \varphi \tag{3.69}$$

Now,

$$v_x = \frac{dx}{dt} = v \cos \varphi$$

Hence,

$$dx = v \cos \varphi\, dt = -\frac{v^2\, d\varphi}{g}$$

where we have used equation (3.69).

Similarly,

$$v_y = \frac{dy}{dt} = v \sin \varphi$$

Hence,

$$dy = v \sin \varphi\, dt = v \tan \varphi \cos \varphi\, dt = -\frac{v^2 \tan \varphi}{\rho}\, d\varphi$$

$$= -\frac{v_x^2 \tan \varphi}{g \cos^2 \varphi}\, d\varphi = -\frac{v_x^2}{g}\, d(\tfrac{1}{2} \tan^2 \varphi)$$

Let the height of the projectile be h, i.e. $y_{max} = h$. Thus, as y varies from 0 to h, angle φ varies from α to 0. Hence

$$\int_0^h \frac{dy}{v_x^2} = -\frac{1}{g}\int_\alpha^0 d(\tfrac{1}{2}\tan^2 \varphi) = \frac{\tan^2 \alpha}{2g} \tag{3.70}$$

During the downward motion of the trajectory the height will vary from h to 0 and angle φ will change from 0 to φ^*, where φ^* is the angle made by the velocity vector with the horizontal as it strikes the ground ($y = 0$). Hence, for this change,

$$\int_h^0 \frac{dy}{v_x^2} = -\frac{1}{g}\int_0^{\varphi^*} d(\tfrac{1}{2}\tan^2 \varphi) = -\frac{\tan^2 \varphi^*}{2g}$$

or

$$\int_0^h \frac{dy}{v_x^2} = \frac{\tan^2 \varphi^*}{2g} \tag{3.71}$$

Due to the air resistance, v_x, the horizontal component of the velocity is smaller in the later part of the trajectory than in the earlier one. Hence, the value of the integral in equation (3.70) is less than that in

equation (3.71). We have, therefore,

$$\tan \varphi^* > \tan \alpha \quad \text{or} \quad \varphi^* > \alpha$$

Thus, the curvature of the trajectory gets reduced or radius ρ of curvature increases as the projectile progresses and the path is, as shown in Fig. 3.3b, quite different from the parabolic path when the air resistance was neglected.

It is obvious from the equations of motion that the velocity with which the projectile falls on the ground is less than the velocity with which it was projected. This reduction in velocity can be obtained from equation (3.67). Thus, by multiplying equation (3.67) by v, we get

$$v \frac{dv}{dt} = -\frac{k}{m} v^2 - gv \sin \varphi$$

or

$$v\, dv = d(\tfrac{1}{2}v^2) = -\frac{k}{m} v^2\, dt - g\, dy \tag{3.72}$$

since $dy = v \sin \varphi\, dt$. Integrating equation (3.72) from 0 to T, the time of flight during which y changes from 0 through h to 0 again, we get

$$\int_{v_0}^{v^*} d(\tfrac{1}{2}v^2) = -\frac{k}{m} \int_0^T v^2\, dt - g \int_0^T dy$$

where v^* is the velocity of the projectile when it strikes the ground. This gives

$$\frac{v^{*2} - v_0^2}{2} = -\frac{k}{m} \int_0^T v^2\, dt < 0, \text{ since} \int_0^T dy = 0$$

Hence,

$$v^* < v_0$$

3.4 MOTION OF A CHARGED PARTICLE IN ELECTROMAGNETIC FIELD

The force acting on a particle carrying charge e ($e > 0$) in an electromagnetic field is given by Lorentz equation

$$\mathbf{F} = e\mathbf{E} + e\mathbf{v} \times \mathbf{B} \tag{3.73}$$

where $\mathbf{E}$ is the electric intensity, $\mathbf{B}$ the magnetic induction and $\mathbf{v}$ is the velocity of the particle.

In case of the electrostatic field, i.e., when $\mathbf{B} = 0$, $\nabla \times \mathbf{E} = 0$ and the field is described by the scalar electric potential $\Phi(\mathbf{r})$ defined by

$$\mathbf{E}(\mathbf{r}) = -\nabla\Phi(\mathbf{r}) \tag{3.74}$$

The force on the charged particle in the magnetostatic field, i.e., when $\mathbf{E} = 0$ depends upon the charge and velocity of the particle and is given by

$$\mathbf{F}_m = e\mathbf{v} \times \mathbf{B} \tag{3.75}$$

Since div $\mathbf{B} = 0$, we have seen earlier that the magnetic induction can be expressed through a vector potential $\mathbf{A}$ defined by

$$\mathbf{B} = \text{curl } \mathbf{A} = \nabla \times \mathbf{A} \tag{3.76}$$

If the particle is moving in a uniform magnetic field, it can easily be

verified that the vector potential is given by

$$\mathbf{A} = -\tfrac{1}{2}\mathbf{r} \times \mathbf{B} \tag{3.77}$$

where $\mathbf{r}$ is the position vector of the particle.

The equation of motion of a particle of mass m under the action of Lorentz force is

$$m\frac{d\mathbf{v}}{dt} = e\mathbf{E} + e\mathbf{v} \times \mathbf{B} \tag{3.78}$$

We assume that the velocity of the particle is small as compared to that of light and hence relativistic variation of mass is neglected.

We shall consider the motion in some special cases.

(a) Motion in a Constant Electric Field

The electrostatic force on a particle of mass m and charge e is $e\mathbf{E}$, and the equation of motion is

$$\mathbf{F} = m\frac{d\mathbf{v}}{dt} = e\mathbf{E} \tag{3.79}$$

or by introducing the scalar potential, equation (3.79) becomes

$$m\frac{d\mathbf{v}}{dt} = -e\nabla\Phi \tag{3.80}$$

Let us take the dot product of both the sides with velocity $\mathbf{v}$. Then

$$m\frac{d\mathbf{v}}{dt}\cdot\mathbf{v} = -e\,\nabla\Phi\cdot\mathbf{v}$$

or

$$\frac{d}{dt}\left(\tfrac{1}{2}mv^2\right) = -e\left(\frac{\partial\Phi}{\partial x}\frac{dx}{dt} + \frac{\partial\Phi}{\partial y}\frac{dy}{dt} + \frac{\partial\Phi}{\partial z}\frac{dz}{dt}\right)$$

$$= -e\frac{d\Phi}{dt}$$

Thus, we get

$$\frac{d}{dt}\left(\tfrac{1}{2}mv^2 + e\Phi\right) = 0$$

or

$$\tfrac{1}{2}mv^2 + e\Phi = \text{const} \tag{3.81}$$

This relation clearly shows conservation of energy which can be stated as "*the sum of kinetic energy $\frac{1}{2}mv^2$ and potential energy $e\Phi$ is a constant.*"

If the particle is initially at rest and falls through the potential difference of V volt, its initial potential energy eV is converted into kinetic energy. Hence,

$$\tfrac{1}{2}mv^2 = eV$$

or

$$v = \sqrt{\frac{2eV}{m}}$$

This gives the velocity acquired by the particle in the direction of the field. For an electron having mass $m = 9.11 \times 10^{-31}$ kg and charge $e = 1.60 \times 10^{-19}$ coul. the velocity is

$$v \approx 6 \times 10^5\sqrt{V}\ \text{m/s} = 600\sqrt{V}\ \text{km/s}$$

Thus electron falling through one volt potential difference gains a velocity of about 600 km/s. The kinetic energy acquired by the electron is measured in terms of electron-volt which is defined by

$$1 \text{ eV} = 1.6 \times 10^{-19} \text{ J}$$

If we take a constant electric field along the y-axis, i.e., $\mathbf{E} = (0, E, 0)$, then the equations of motion in the cartesian coordinate system are

$$m\ddot{x} = 0 = m\ddot{z}, \quad m\ddot{y} = eE \tag{3.82}$$

If the particle is initially at rest, it will move along the y-direction only. The solutions of equations (3.82) are

$$x = x_0, \quad y = y_0 + \tfrac{1}{2}at^2, \quad z = z_0 \tag{3.83}$$

where (x_0, y_0, z_0) is the original position of the particle, and $a = \dfrac{eE}{m}$ the acceleration due to the electrostatic field along the y-axis.

(b) Motion in a Constant Magnetic Field

Now, the equation of motion of the charged particle is

$$m\frac{d\mathbf{v}}{dt} = e\mathbf{v} \times \mathbf{B} \tag{3.84}$$

Taking the dot product of both sides of equation (3.84) with velocity $\mathbf{v}$, we get

$$m\frac{d\mathbf{v}}{dt}\cdot\mathbf{v} = e\mathbf{v}\cdot(\mathbf{v} \times \mathbf{B}) \tag{3.85}$$

i.e.

$$\frac{d}{dt}(\tfrac{1}{2}mv^2) = 0$$

since the right-hand side is a scalar triple product and $\mathbf{v}\cdot(\mathbf{v}\times\mathbf{B}) = (\mathbf{v}\times\mathbf{v})\cdot\mathbf{B} = 0$. Thus, the kinetic energy of the particle remains unchanged. Hence,

$$\tfrac{1}{2}mv^2 = \text{const} \tag{3.86}$$

Thus, speed v of the particle is unchanged in the magnetic field and field $\mathbf{B}$ changes only the direction of its velocity. The displacement $d\mathbf{r}$ of the particle in the magnetic field is, therefore, always perpendicular to the magnetic force on the particle and hence no work is done during this displacement.

Now consider the case of uniform magnetic field given by $\mathbf{B}$ = constant. Let us split velocity $\mathbf{v}$ into two components $\mathbf{v}_\perp$ which is perpendicular and $\mathbf{v}_\parallel$ which is parallel to the magnetic induction $\mathbf{B}$ (Fig. 3.4a). Thus,

$$\mathbf{v} = \mathbf{v}_\perp + \mathbf{v}_\parallel \tag{3.87}$$

The parallel and perpendicular components of magnetic force are

$$\mathbf{F}_{m_\parallel} = e\mathbf{v}_\parallel \times \mathbf{B} = 0 \tag{3.88}$$

and

$$\mathbf{F}_{m_\perp} = e\mathbf{v}_\perp \times \mathbf{B} \tag{3.89}$$

Thus, from equation (3.88), we have

$$\mathbf{F}_{m_\parallel} = m\frac{d\mathbf{v}_\parallel}{dt} = 0 \quad \text{or} \quad \mathbf{v}_\parallel = \text{const} \tag{3.90}$$

Thus, the velocity of the particle parallel to the field is unaffected. From equation (3.89), we have

$$m\frac{d\mathbf{v}_\perp}{dt} = e\mathbf{v}_\perp \times \mathbf{B} \tag{3.91}$$

But, from equation (3.87), we have

$$v^2 = v_\perp^2 + v_\parallel^2$$

and since $v_\parallel^2$ is constant by equation (3.90) and v^2 is constant by equation (3.81)

$$v_\perp^2 = \text{const} \tag{3.92}$$

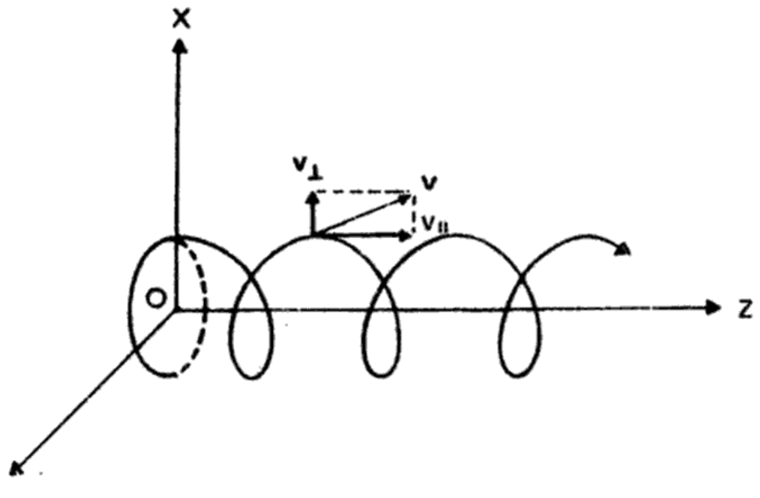


Fig. 3.4a Spiral motion of a charged particle along the constant magnetic induction **B** taken along the z-axis

Thus, the magnitude of perpendicular component is unchanged but its direction changes according to equation (3.91). We are considering motion in a constant magnetic induction and $\mathbf{v}_\perp$ is perpendicular to **B**. Hence $v_\perp B$ is a constant quantity. Thus the particle in a constant magnetic field moves in such a way that the force $ev_\perp B$ and acceleration $\frac{eB}{m}v_\perp$ are always directed perpendicular to velocity. This gives a circular motion in the plane perpendicular to **B** and the force is the centripetal force (Fig. 3.4b). If the circle has radius ρ then equation (3.89) can be written as

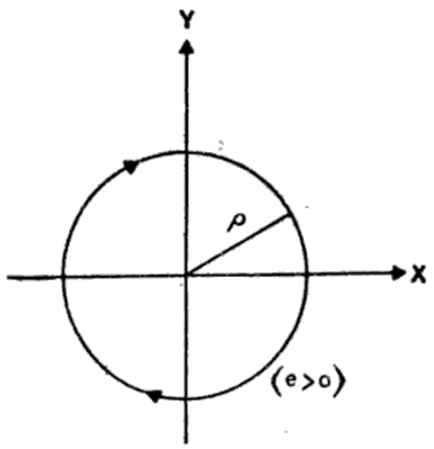


Fig. 3.4b Projection of the spiral motion on the xy-plane

$$\frac{mv_\perp^2}{\rho} = ev_\perp B$$

or

$$v_\perp = \frac{eB}{m}\rho$$

The angular frequency of the circular

motion is given by

$$\omega = \frac{2\pi}{\text{period}} = \frac{2\pi}{(2\pi\rho/v_\perp)} = \frac{v_\perp}{\rho} = \frac{eB}{m}$$

and is known as *cyclotron frequency*. The radius of the circular orbit is given by

$$\rho = \frac{v_\perp}{\omega} = \frac{mv_\perp}{eB}$$

In order to find the orbit of the particle, let magnetic induction $\mathbf{B}$ be taken along the z-axis so that $\mathbf{B} = \mathbf{k}B$. In the cartesian coordinates, equation of motion (3.84) now becomes

$$\frac{d\mathbf{v}}{dt} = \omega\mathbf{v} \times \mathbf{k} \tag{3.93}$$

where $\omega = \dfrac{eB}{m}$ Or in component form

$$\frac{dv_x}{dt} = \omega v_y, \ \frac{dv_y}{dt} = -\ \omega v_x, \ \frac{dv_z}{dt} = 0 \tag{3.94}$$

The first two equations in (3.94) are coupled equations, i.e., acceleration along the x-direction depends on velocity along the y-direction and vice versa. The third of equations (3.94) expresses the fact that velocity along the z-direction, i.e., along $\mathbf{B}$ is constant and its solution is

$$z = z_0 + v_\parallel t \tag{3.95}$$

where $\dot{z} = v_z = v_\parallel =$ constant. Thus, the particle moving from z_0 along the z-axis will be unaffected by the magnetic field and will move with constant velocity.

The two coupled equations can be combined together by multiplying the second equation by $i = \sqrt{-1}$ and then adding them together. Thus,

$$\frac{d}{dt}(v_x + iv_y) = -\ i\omega(v_x + iv_y) \tag{3.96}$$

The solution of this equation is clearly

$$v_x + iv_y = Ce^{-i\omega t} \tag{3.97}$$

Second integration gives

$$x + iy = C_1 e^{-i\omega t} + C_2$$

where C_1 and C_2 are constants of integration to be determined from initial position and velocity. We choose the constants in the form

$$C_1 = Ae^{-i\alpha} \text{ and } C_2 = x_0 + iy_0$$

Then, the solution of the coupled equation is

$$x + iy = Ae^{-i(\omega t + \alpha)} + x_0 + iy_0 \tag{3.98}$$

The x- and y-coordinates of the particle are obviously real and are obtained from equation (3.98). Therefore,

$$\left.\begin{aligned} x &= \text{Re}\,[Ae^{-i(\omega t+\alpha)} + x_0 + iy_0] = A\cos(\omega t + \alpha) + x_0 \\ y &= \text{Im}\,[Ae^{-i(\omega t+\alpha)} + x_0 + iy_0] = -A\sin(\omega t + \alpha) + y_0 \end{aligned}\right\} \tag{3.99}$$

Equations (3.99) together with equation (3.95) give the trajectory of the particle in a constant magnetic field.

From equation (3.99) we have

$$(x - x_0)^2 + (y - y_0)^2 = A^2 \tag{3.100}$$

which is an equation of a circle in the xy-plane with (x_0, y_0) as the origin. If $v_{\parallel} = 0$, then the particle will move in a circular orbit with angular velocity vector

$$\boldsymbol{\omega} = -\frac{e\mathbf{B}}{m} = -\frac{eB}{m}\mathbf{k} = -\omega\mathbf{k} \tag{3.101}$$

The radius $A(=\rho)$ is obtained from equation (3.99) and remembering that $\dot{x}^2 + \dot{y}^2 + \dot{z}^2 = v_{\perp}^2$. Thus,

$$\dot{x}^2 + \dot{y}^2 = v_{\perp}^2 = A^2\omega^2$$

or

$$A = \frac{v_{\perp}}{\omega} = \rho \tag{3.102}$$

as is already obtained above.

(c) Motion in Crossed Fields

Consider now the motion of a charged particle when constant electric and magnetic fields are simultaneously acting over a region of space. The equation of motion of the particle is

$$m\frac{d\mathbf{v}}{dt} = e\mathbf{E} + e\mathbf{v} \times \mathbf{B} \tag{3.103}$$

Let the fields be perpendicular to each other, i.e. $\mathbf{E} \perp \mathbf{B}$ (Fig. 3.5). We search for the solution of equation (3.103) in the form

$$\mathbf{v} = \mathbf{v}' + \frac{1}{B^2}\mathbf{E} \times \mathbf{B} = \mathbf{v}' + \mathbf{v}_d \tag{3.104}$$

where we have put $\mathbf{v}_d = \frac{1}{B^2}(\mathbf{E} \times \mathbf{B})$.

Substituting $\mathbf{v}$ from equation (3.104) in equation (3.103) and noting that

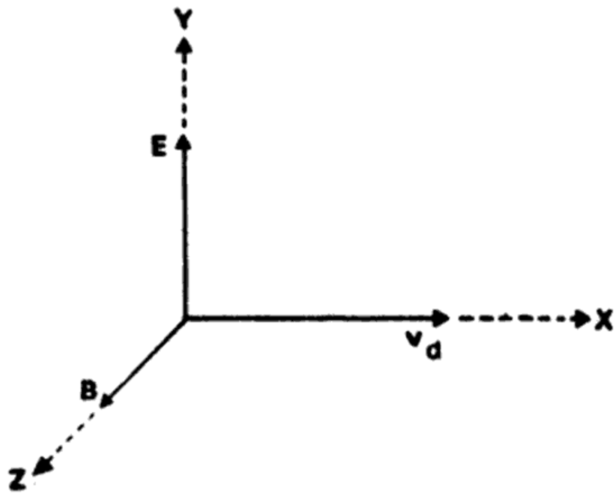


Fig. 3.5 Crossed fields: $\mathbf{E} \perp \mathbf{B}$

E and **B** are constants, we get

$$m\frac{d\mathbf{v}'}{dt} = e\mathbf{E} + e\mathbf{v}' \times \mathbf{B} + \frac{e}{B^2}(\mathbf{E} \times \mathbf{B}) \times \mathbf{B} \tag{3.105}$$

since
$$\frac{d}{dt}\frac{\mathbf{E} \times \mathbf{B}}{B^2} = 0.$$

But
$$\frac{e}{B^2}(\mathbf{E} \times \mathbf{B}) \times \mathbf{B} = \frac{e}{B^2}(\mathbf{E} \cdot \mathbf{B})\,\mathbf{B} - e\mathbf{E} = -\,e\mathbf{E}$$

because the **E** and **B** fields are mutually perpendicular. Hence, equation (3.105) becomes

$$m\frac{d\mathbf{v}'}{dt} = e\mathbf{v}' \times \mathbf{B}$$

Thus, we have

$$m\frac{d\mathbf{v}_d}{dt} = 0 \text{ and } m\frac{d\mathbf{v}'}{dt} = e\mathbf{v}' \times \mathbf{B} \tag{3.106}$$

The first of equations (3.106) suggests that $\mathbf{v}_d$ is constant. Velocity $\mathbf{v}_d$ is known as the *drift velocity*. The second equation of (3.106) is similar to equation (3.91). However, it does not describe circular motion since, in this case, the electrostatic field modifies the motion. The magnitude of the drift velocity is

$$v_d = \frac{|\mathbf{E} \times \mathbf{B}|}{B^2} = \frac{E}{B} \tag{3.107}$$

and is fixed by the ratio of electric intensity and magnetic induction.

Let us now obtain the equation for the trajectory of a particle moving in crossed constant fields by solving the equation of motion (3.103) in cartesian coordinates. Consider a general case when **E** and **B** are not perpendicular, but make some angle (Fig. 3.6). We can take **B** along the

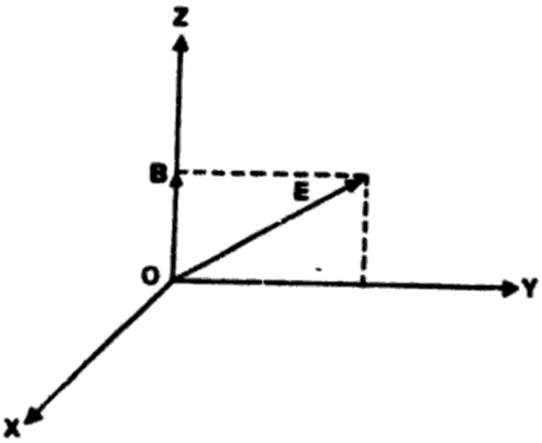


Fig. 3.6 Crossed fields

z-axis and **E** in the yz-plane. Let initially, i.e. at $t = 0$, the particle be at origin and let its initial velocity be $\mathbf{v}_0 = (v_{0x}, v_{0y}, v_{0z})$. In this case, equations (3.103) in the component form are

$$\frac{dv_x}{dt} = \omega v_y \tag{3.108a}$$

$$\frac{dv_y}{dt} = a_y - \omega v_x \tag{3.108b}$$

$$\frac{dv_z}{dt} = a_z \tag{3.108c}$$

where we have put $a_y = \frac{eE_y}{m}$, $a_z = \frac{eE_z}{m}$ and $\omega = \frac{eB}{m}$.

Equation (3.108c) has solutions

$$v_z = v_{0z} + a_z t \text{ and } z = v_{0z}t + \tfrac{1}{2}a_z t^2 \tag{3.109}$$

Coupled equations (3.108a) and (3.108b) can be solved by differentiating any one of them with respect to time and substituting the time derivative of velocity from the other.

$$\frac{d^2v_x}{dt^2} = \omega\frac{dv_y}{dt} = a_y\omega - \omega^2 v_x \tag{3.110}$$

and

$$\frac{d^2v_y}{dt^2} = -\omega\frac{dv_x}{dt} = -\omega^2 v_y \tag{3.111}$$

Equation (3.111) is similar to the equation of a simple harmonic oscillator and has solution

$$v_y = A\sin\omega t \tag{3.112}$$

The other constant, i.e., the phase will be taken as zero.

Substituting solution (3.112) in equation (3.108a) and integrating it with respeci to time, we get

$$v_x - v_{0x} = \omega A\int_0^t \sin\omega t\, dt = A(1-\cos\omega t) \tag{3.113}$$

Further integration of equations (3.113) and (3.112) yields

$$x = \int_0^t v_x\, dt = v_{0x}t + \frac{A}{\omega}(\omega t - \sin\omega t) \tag{3.114}$$

and

$$y = \int_0^t v_y\, dt = \frac{A}{\omega}(1-\cos\omega t) \tag{3.115}$$

These are the parametric equations of a cycloid. Some forms of trajectories are illustrated in Fig. 3.7. Displacement y becomes zero whenever $\omega t = 2\pi n$, where n is an integer. Under this condition, $v_x = v_{0x}$ and $v_y = 0$.

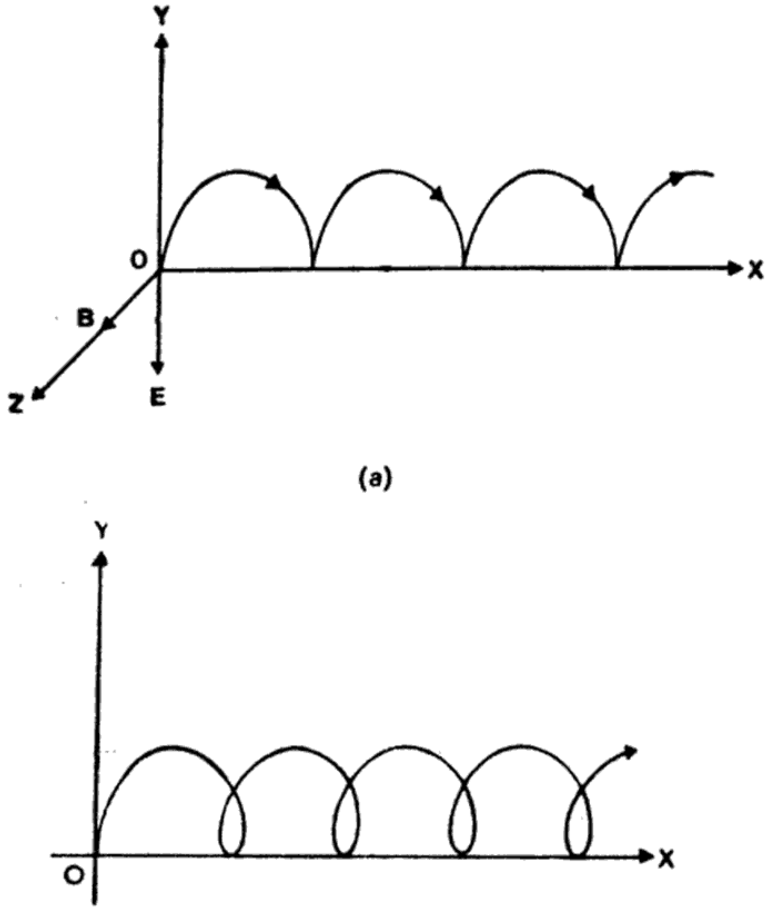
Equation (3.109) shows that the particle moves in the direction of the z-axis because of initial velocity component and the component of electric field along the z-axis.

Consider crossed fields $\mathbf{E} \perp \mathbf{B}$, i.e., let $E_z = 0$ Now, the particle moves along the z-axis as given by

$$z = v_{0z}t \tag{3.116}$$

Let $v_{0x} = 0$, $v_{0y} = 0$. In this case the trajectory of the particle is a cycloid and constant A can be obtained from equations (3.112) and (3.108b) by taking values at $t = 0$. Thus, from

$$\frac{dv_y}{dt} = a_y - \omega v_x = A\omega\cos\omega t$$



(b)
Fig. 3.7b Different paths of a charged particle in crossed fields

at $t = 0$, we get

$$A = \frac{a_y}{\omega} - v_{0x} = \frac{E_y}{B} - v_{0x} \tag{3.117}$$

Since $v_{0x} = 0$ and $E_y = E$, we have

$$A = \frac{E}{B}$$

The case of crossed electric and magnetic fields is of great practical application. If initially, the particle is moving along the x-direction with velocity v_{0x}, so that we have at $t = 0$, $x = y = z = 0$, $v_{0y} = 0 = v_{0z}$, and $v_{0x} \neq 0$, the solution (3.112) in this case is

$$v_y = \left(\frac{E_y}{B} - v_{0x}\right) \sin \omega t \tag{3.118}$$

If $v_{0x} = \dfrac{E_y}{B}$, the particle will not be deflected along the y-axis. Thus, the particles with velocity (E_y/B) will go undeviated and are said to be filtered out. The perpendicular combination of electric and magnetic fields thus works as a velocity filter for charged particles, and particles of desired velocity can be obtained by choosing suitable values of the electric intensity **E** and magnetic induction **B**.

3.5 MECHANICS OF SYSTEMS OF PARTICLES

We now extend the considerations of the previous section to a system of N particles. Let $\mathbf{F}^{ext}$ be the extesnal force acting on the system. Consider the ith particle of the system having mass m_i and let its position vector with respect to a fixed origin be $\mathbf{r}_i$ (Fig. 3.8). This particle will be

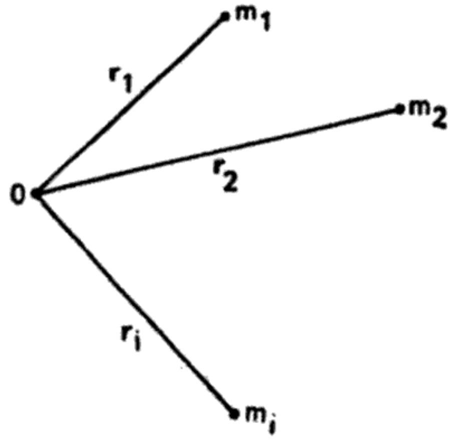


Fig. 3.8 System of particles

acted upon by two forces—external force $\mathbf{F}_i^{ext}$ and internal force $\mathbf{F}_i^{int}$ due to the interaction between the particles.

Total mass M of the system is given by

$$M = \sum_{i=1}^{N} m_i \tag{3.119}$$

The *centre of mass* of the system is defined as *a point whose position vector is given by*

$$\mathbf{R} = \frac{\sum m_i \mathbf{r}_i}{\sum m_i} = \frac{1}{M} \sum m_i \mathbf{r}_i \tag{3.120}$$

Let $\mathbf{F}_{ij}^{int}$ be the internal force acting on the ith particle as a result of its interaction with the jth particle. Then, the total internal force acting on the ith particle is

$$\mathbf{F}_i^{int} = \sum_j \mathbf{F}_{ij}^{int},\ i \neq j$$

Then, the equation of motion of the ith particle can be written as

$$\dot{\mathbf{p}}_i = \mathbf{F}_i^{ext} + \sum_j \mathbf{F}_{ij}^{int} \tag{3.121}$$

The first term on the right-hand side of equation (3.121) represents the external force on the ith particle, while the second term is a vector sum of all the internal forces due to the interaction of the remaining $N-1$ particles with the ith particle. In this summation we shall take $\mathbf{F}_{ii}^{int} = 0$, i.e., the force of self-interaction is zero.

Now, according to Newton's third law of motion

$$\mathbf{F}_{ij}^{int} = -\mathbf{F}_{ji}^{int} \tag{3.122}$$

i.e., the mutual interaction forces between the ith and jth particle are equal and opposite. Illustrations of such force-fields are the gravitational and

Thus, if $\mathbf{F}^{ext} = 0$, $\dot{\mathbf{P}} = 0$ or $\mathbf{P} = \text{const}$ (3.128)

(a) Angular Momentum of the System

The total angular momentum of the system about any point will be equal to the vector sum of the angular momenta of individual particles. Let $\mathbf{l}_i$ represent the angular momentum of the ith particle about some point. Then,

$$\mathbf{l}_i = \mathbf{r}_i \times \mathbf{p}_i \tag{3.129}$$

where $\mathbf{r}_i$ is the position vector of the ith particle from the given point.

Hence the total angular momentum of the system is obtained by taking vector sum of individual momenta.

$$\mathbf{L} = \sum_i \mathbf{l}_i = \sum_i \mathbf{r}_i \times \mathbf{p}_i \tag{3.130}$$

Let $\mathbf{N}$ be the total torque acting on the system. Then,

$$\mathbf{N} = \frac{d\mathbf{L}}{dt} = \frac{d}{dt}\sum_i \mathbf{r}_i \times \mathbf{p}_i$$

$$= \sum_i \dot{\mathbf{r}}_i \times \mathbf{p}_i + \sum_i \mathbf{r}_i \times \dot{\mathbf{p}}_i$$

But, $$\sum_i \dot{\mathbf{r}}_i \times \mathbf{p}_i = \sum_i \dot{\mathbf{r}}_i \times m_i\dot{\mathbf{r}}_i = \sum_i m_i\dot{\mathbf{r}}_i \times \dot{\mathbf{r}}_i = 0$$

Further, $$\sum_i \mathbf{r}_i \times \dot{\mathbf{p}}_i = \sum_i \mathbf{r}_i \times (\mathbf{F}_i^{ext} + \sum_j{}' \mathbf{F}_{ij}^{int})$$

$$= \sum \mathbf{r}_i \times \mathbf{F}_i^{ext}$$

Since $$\sum_i \sum_j{}' \mathbf{r}_i \times \mathbf{F}_{ij}^{int} = +\sum_j \sum_i{}' \mathbf{r}_j \times \mathbf{F}_{ji}^{int}$$

by the same argument as given above and

$$\sum_{ij}{}' \mathbf{r}_i \times \mathbf{F}_{ij}^{int} = \tfrac{1}{2}\sum_{ij}{}' [\mathbf{r}_i \times \mathbf{F}_{ij}^{int} + \mathbf{r}_j \times \mathbf{F}_{ji}^{int}]$$

$$= \tfrac{1}{2}\sum_{ij}{}' [(\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{F}_{ij}^{int}]$$

$$= \tfrac{1}{2}\sum_{ij}{}' \mathbf{r}_{ij} \times \mathbf{F}_{ij}^{int}$$

Now, $\mathbf{F}_{ij}^{int}$ is also proportional to $(\mathbf{r}_i - \mathbf{r}_j) \equiv \mathbf{r}_{ij}$, since we are considering the forces of action and reaction only. Hence, the right-hand side of the above equation will be zero. Then, we have

$$\sum_{ij}{}' \mathbf{r}_i \times \mathbf{F}_{ij}^{int} = 0$$

This is true in the case of internal forces acting along the line joining the two particles but not in the case of forces acting on the moving charged particles.

Thus, total torque $\mathbf{N}$ is given by

$$\mathbf{N} = \frac{d\mathbf{L}}{dt} = \sum_i \mathbf{r}_i \times \mathbf{F}_i^{ext} = \sum_i \mathbf{N}_i \tag{3.131}$$

Equation (3.131) shows that the total torque on the system is equal to the vector sum of the torques acting on the individual particles.

If $\mathbf{N} = 0, \dfrac{d\mathbf{L}}{dt} = 0$ or $\mathbf{L} = \text{const}$ (3.132)

Thus, if the total external torque acting on the system is zero, then the total angular momentum of the system is conserved.

Now, we prove that *the angular momentum of the system about a fixed point is equal to the sum of the angular momentum of the total mass concentrated at the centre of mass about that point and the angular momentum of the system about its centre of mass.*

Let m_i and $\mathbf{r}_i$ denote the mass and the position vector of the ith particle with reference to point O (Fig. 3.9). Let $\mathbf{r}_i'$ be the position

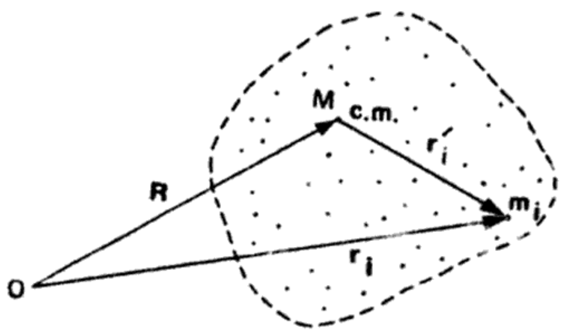


Fig. 3.9 System of particles with total mass M and centre of mass at $\mathbf{R}$

vector of the ith particle with reference to the centre of mass of the system. Then

$$\mathbf{r}_i = \mathbf{R} + \mathbf{r}_i' \tag{3.133}$$

The corresponding equation relating the velocities can be written as

$$\dot{\mathbf{r}}_i = \dot{\mathbf{R}} + \dot{\mathbf{r}}_i'$$

or

$$\mathbf{v}_i = \mathbf{V} + \mathbf{v}_i' \tag{3.134}$$

where $\mathbf{V}$ is the velocity of the centre of mass of the system and $\mathbf{v}_i'$ is the velocity of the ith particle with reference to the centre of mass of the system.

The total angular momentum is, then, given by

$$\begin{aligned}\mathbf{L} &= \sum \mathbf{r}_i \times \mathbf{p}_i \\ &= \sum_i (\mathbf{R} + \mathbf{r}_i') \times m_i(\dot{\mathbf{R}} + \dot{\mathbf{r}}_i') \\ &= \sum_i m_i \mathbf{R} \times \dot{\mathbf{R}} + \sum_i \mathbf{R} \times m_i \dot{\mathbf{r}}_i' + \sum_i m_i \mathbf{r}_i' \times \dot{\mathbf{R}} + \sum_i m_i \mathbf{r}_i' \times \dot{\mathbf{r}}_i'\end{aligned} \tag{3.135}$$

The second term on the right-hand side of equation (3.135) can be written as

$$\begin{aligned}\sum_i \mathbf{R} \times m_i \dot{\mathbf{r}}_i' &= \mathbf{R} \times \sum_i m_i \dot{\mathbf{r}}_i' \\ &= \mathbf{R} \times \frac{d}{dt} \sum_i m_i \mathbf{r}_i'\end{aligned}$$

But, $\sum_i m_i \mathbf{r}_i' = 0$, since distances $\mathbf{r}_i'$ are measured with respect to the centre of mass of the system.

The third term on the right-hand side of equation (3.135) also vanishes for the same reason. Then, we are left with

$$\begin{aligned} \mathbf{L} &= \sum_i m_i \mathbf{R} \times \dot{\mathbf{R}} + \sum_i m_i \mathbf{r}_i' \times \dot{\mathbf{r}}_i' \\ &= \mathbf{R} \times \dot{\mathbf{R}} \sum_i m_i + \sum_i \mathbf{r}_i' \times m_i \dot{\mathbf{r}}_i' \\ &= \mathbf{R} \times \dot{\mathbf{R}} M + \sum_i \mathbf{r}_i' \times \mathbf{p}_i' \end{aligned}$$

or

$$\mathbf{L} = \mathbf{R} \times \mathbf{P} + \mathbf{L}' \tag{3.136}$$

where $\mathbf{P} = M\dot{\mathbf{R}}$ represents the linear momentum of the centre of mass and hence $\mathbf{R} \times \mathbf{P}$ represents the angular momentum of the total mass concentrated at the centre of mass about point O. The term $\mathbf{L}' = \sum_i \mathbf{r}_i' \times \mathbf{p}_i'$ represents the angular momentum of the system about its centre of mass. This proves the statement.

(b) Energy of the System

In order to find the energy of the system, let us find the work done by all the forces—external as well as internal—in moving the system from initial configuration 1 to final configuration 2. The total work done in moving the system is equal to the sum of the work done in moving all the particles from configuration 1 to configuration 2. Thus,

$$\begin{aligned} W_{12} &= \sum_i \int_1^2 \mathbf{F}_i \cdot d\mathbf{r}_i \\ &= \sum_i \int_1^2 \mathbf{F}_i^{\text{ext}} \cdot d\mathbf{r}_i + \sum_{ij}{}' \int_1^2 \mathbf{F}_{ij}^{\text{int}} \cdot d\mathbf{r}_i \end{aligned} \tag{3.137}$$

If the internal and the external forces are conservative, then they can be expressed in terms of corresponding potential energies. Thus, total force $\mathbf{F}_i$ on the ith particle can be written as

$$\mathbf{F}_i = \mathbf{F}_i^{\text{ext}} + \sum_j \mathbf{F}_{ij}^{\text{int}} = -\nabla_i V_i \tag{3.138a}$$

where the potential energy

$$V_i = V_i^{\text{ext}} + V_i^{\text{int}} \tag{3.138b}$$

is the sum of potential energy functions of the external and internal forces. In equation (3.138a), symbol ∇_i is

$$\nabla_i = \sum_i \hat{\mathbf{e}}_i \frac{\partial}{\partial x_i} \tag{3.139}$$

and represents the gradient operator performing differentiation with respect to x_i, components of position vector $\mathbf{r}_i$ of the ith particle. The operator can be written separately as

$$\left.\begin{aligned} \mathbf{F}_i^{\text{ext}} &= -\nabla_i V_i^{\text{ext}} \\ \text{and} \qquad \mathbf{F}_{ij}^{\text{int}} &= -\nabla_{ij} V_{ij}^{\text{int}} \end{aligned}\right\} \tag{3.140}$$

Quantity V_{ij}^{int} is the potential energy arising due to internal forces $\mathbf{F}_{ij}^{\text{int}}$ and $\nabla_{ij} = \sum_i \hat{\mathbf{e}}_i \frac{\partial}{\partial(x_i - x_j)}$. From this, it will be clear that $\nabla_{ij} = -\nabla_{ji}$.

Now the potential energy of the ith particle arising due to internal forces is given by

$$V_i^{\text{int}} = \sum_j V_{ij}^{\text{int}}$$

Hence, the total potential energy due to internal forces is

$$V^{\text{int}} = \sum_i V_i^{\text{int}} = \sum_{\substack{ij\\ i<j}} V_{ij}^{\text{int}} \tag{3.141}$$

We have to take $V_{ii}^{\text{int}} = 0$, to have $\mathbf{F}_{ii}^{\text{int}} = 0$. Condition $i < j$ is necessary because otherwise each term will be taken twice in summing over i and j.

Potential energy V_{ij}^{int} depends upon the relative positions of the two particles, i.e. $V_{ij}^{\text{int}} = V_{ij}^{\text{int}}(\mathbf{r}_{ij})$. Then, by Newton's third law, we have

$$\begin{aligned}\mathbf{F}_{ij}^{\text{int}} &= -\nabla_{ij} V_{ij}^{\text{int}} \\ &= -\mathbf{F}_{ji}^{\text{int}} = +\nabla_{ji} V_{ji}^{\text{int}} = -\nabla_{ij} V_{ji}^{\text{int}}\end{aligned} \tag{3.142}$$

From equation (3.142), we find that $V_{ij}^{\text{int}} = V_{ji}^{\text{int}}$ and hence

$$V^{\text{int}} = \frac{1}{2} \sum_{ij} V_{ij}^{\text{int}} \tag{3.143}$$

Factor $\frac{1}{2}$ occurring on the right-hand side of equation (3.143) is due to the fact that each term is being taken twice while summing over i and j. The same has been incorporated in equation (3.141) by writing $i<j$. This condition avoids the duplication of terms.

The work done by the external force is given by

$$\begin{aligned}\sum_i \int_1^2 \mathbf{F}_i^{\text{ext}} \cdot d\mathbf{r}_i &= -\sum_i \int_1^2 \nabla_i V_i^{\text{ext}} \cdot d\mathbf{r}_i \\ &= -\sum_i \int_1^2 dV_i^{\text{ext}} \\ &= -\sum_i V_i^{\text{ext}} \Big|_1^2 \\ &= V_1^{\text{ext}} - V_2^{\text{ext}}\end{aligned} \tag{3.144}$$

where V_1^{ext} and V_2^{ext} represent the potential energies of the system arising due to external forces acting on the system in configuration 1 and 2 respectively.

The work done by the internal forces is given by

$$\sum_{ij}{}' \int_1^2 \mathbf{F}_{ij}^{\text{int}} \cdot d\mathbf{r}_i$$

But, $$\sum_{ij}{}' \mathbf{F}_{ij}^{\text{int}} \cdot d\mathbf{r}_i = \sum_{ij}{}' \mathbf{F}_{ji}^{\text{int}} \cdot d\mathbf{r}_j = -\sum_{ij}{}' \mathbf{F}_{ij}^{\text{int}} \cdot d\mathbf{r}_j$$

Hence, $$\sum_{ij}{}' \mathbf{F}_{ij}^{\text{int}}\cdot d\mathbf{r}_i = \tfrac{1}{2}\sum_{ij}{}' \mathbf{F}_{ij}^{\text{int}}\cdot(d\mathbf{r}_i - d\mathbf{r}_j)$$
$$= \tfrac{1}{2}\sum_{ij} \mathbf{F}_{ij}^{\text{int}}\cdot d\mathbf{r}_{ij}$$
where $d\mathbf{r}_{ij} = d\mathbf{r}_i - d\mathbf{r}_j$.

Substituting this value, we get
$$\sum_{ij}{}' \int_1^2 \mathbf{F}_{ij}^{\text{int}}\cdot d\mathbf{r}_i = \tfrac{1}{2}\sum_{ij}{}' \int_1^2 \mathbf{F}_{ij}^{\text{int}}\cdot d\mathbf{r}_{ij}$$
$$= -\tfrac{1}{2}\sum_{ij}{}' \int_1^2 \nabla_{ij} V_{ij}^{\text{int}}\cdot d\mathbf{r}_{ij}$$
$$= -\tfrac{1}{2}\sum_{ij}{}' \int_1^2 dV_{ij}^{\text{int}}$$
$$= -\tfrac{1}{2}\sum_{ij}{}' V_{ij}^{\text{int}}\Big|_1^2$$
$$= -V^{\text{int}}\Big|_1^2, \text{ by equation (3.143)}$$
$$= V_1^{\text{int}} - V_2^{\text{int}} \tag{3.145}$$
The total potential energy of the system is then given by
$$V = V^{\text{ext}} + V^{\text{int}}$$
$$= \sum_i V_i^{\text{ext}} + \tfrac{1}{2}\sum_{ij}{}' V_{ij}^{\text{int}}$$
$$= \sum_i V_i^{\text{ext}} + \sum_{\substack{ij\\ i<j}} V_{ij}^{\text{int}}$$
In terms of the total potential energy of the system, the work done is
$$W_{12} = -V\Big|_1^2 = V_1 - V_2 \tag{3.146}$$
Work done W_{12} can be expressed in terms of the difference between the kinetic energy of the system in the initial and final configurations as follows:
$$W_{12} = \sum_i \int_1^2 \mathbf{F}_i\cdot d\mathbf{r}_i$$
$$= \sum_i \int_1^2 \frac{d}{dt}(m_i\mathbf{v}_i)\cdot\frac{d\mathbf{r}_i}{dt}\,dt = \sum_i \int_1^2 m_i\left(\frac{d\mathbf{v}_i}{dt}\cdot\mathbf{v}_i\right)dt$$
$$= \sum_i \int_1^2 m_i(\mathbf{v}_i\cdot d\mathbf{v}_i)$$
$$= \sum_i \int_1^2 d(\tfrac{1}{2}m_i v_i^2)$$
$$= \sum_i T_i\Big|_1^2 = T\Big|_1^2 = T_2 - T_1 \tag{3.147}$$
Combining equations (3.146) and (3.147), we get
$$T_1 + V_1 = T_2 + V_2$$
or $$E_1 = E_2 \tag{3.148}$$
Equation (3.148) states that the total energy of the system is conserved.

This is true only when *all* the forces—internal and external—are derivable from the potential energy functions that do not depend explicitly on time. A system in which all the *forces acting on it are derivable from the potential energy functions is called a conservative system.*

In a conservative system, potential energy V_{ij}^{int} depends entirely upon the separation between the ith and jth particle. If this separation is constant, V_{ij} is also constant and can be taken to be zero. A system in which the distance between any two particles remains constant is called a *rigid body*.

(c) Kinetic Energy of the System

We now express the kinetic energy of the system as a sum of (i) the kinetic energy of a particle at the centre of mass and having mass $M = \sum m_i$ and moving with velocity $\mathbf{V} = \dot{\mathbf{R}}$, with respect to origin O, and (ii) the kinetic energy of the system referred to the centre of mass as the origin. The total kinetic energy of the system is

$$T = \sum T_i = \sum_i \tfrac{1}{2} m_i v_i^2 \tag{3.149}$$

But by equation (3.134), we have, $\mathbf{v}_i = \mathbf{V} + \mathbf{v}_i'$.

Hence,

$$\begin{aligned} v_i^2 &= \mathbf{v}_i \cdot \mathbf{v}_i \\ &= (\mathbf{V} + \mathbf{v}_i') \cdot (\mathbf{V} + \mathbf{v}') \\ &= V^2 + v_i'^2 + 2\mathbf{V} \cdot \mathbf{v}_i' \end{aligned}$$

Substituting this value in equation (3.149), we get

$$\begin{aligned} T &= \sum_i \tfrac{1}{2} m_i [V^2 + v'^2 + 2\mathbf{V} \cdot \mathbf{v}_i'] \\ &= \sum_i \tfrac{1}{2} m_i V^2 + \sum_i \tfrac{1}{2} m_i v_i'^2 + \sum_i \tfrac{1}{2} m_i 2\mathbf{V} \cdot \mathbf{v}_i' \end{aligned} \tag{3.150}$$

The third term on the right-hand side of equation (3.150) can be shown to be zero as follows:

$$\begin{aligned} \sum \tfrac{1}{2} m_i 2\mathbf{V} \cdot \mathbf{v}_i' &= \mathbf{V} \cdot \sum_i m_i \mathbf{v}_i' \\ &= \mathbf{V} \cdot \frac{d}{dt} \sum m_i \mathbf{r}_i' \\ &= 0 \end{aligned}$$

Hence,

$$T = \tfrac{1}{2} M V^2 + T_c \tag{3.151}$$

where $T_c = \sum_i \frac{1}{2} m_i v_i'^2 =$ kinetic energy of the system about the centre of mass of the system.

This proves the statement.

(d) Laws of Conservation

In the previous articles, we have obtained the laws of conservation of momentum and energy of a particle or of a system of particles as a consequence of Newton's laws of motion. These laws of conservation are helpful in analysing the motion of a particle or of a system, particularly when

the nature of the force is not known. The conservation laws relate the momentum or the energy of a system at two different instants and help us to obtain the kinematical relations for the system. These laws are used extensively in studying collisions of particles.

The conservation laws have far wider applicability and are not restricted to Newtonian mechanics alone. The laws of conservation are exact, i.e., they are true in the case of systems with any type of interaction between the particles. In fact, so great is the conviction in the laws of conservation that it led W. Pauli in 1930 to postulate a new particle called *neutrino* to account for the missing energy in the process of beta-decay.

The laws of conservation must obviously have an intimate relationship with the physical nature of space and time. The relation of the laws with symmetry is discussed in the Lagrangian formulation of mechanics.

3.6 MOTION OF A SYSTEM WITH VARIABLE MASS

So far, we have considered the equations of motion and the laws of conservation in such cases when the mass of the system was constant during the motion of the system. We now wish to consider the motion of a system when the mass varies with time. We often come across such systems in nature and also in technology. A drop of water falling through a cloud will gain in mass as it descends. A rocket will lose mass in its flight as a result of the burning of the fuel and the exhaust gas which provides acceleration to the rocket to attain high velocities. We can apply the laws of conservation in the case of such systems to obtain the equations of motion of the system and solve them.

It should be noted that we are not considering the variation of mass of a particle with velocity which is the well-known relativistic effect. The velocities involved in the problems under discussion are very small as compared to the velocity of light. Hence, the discussion that follows is a non-relativistic discussion.

A rocket fired from the earth will always be affected by the gravitational pull of the earth. Other celestial objects are at great distances from the rocket and the effect of such objects on the motion of the rocket can be ignored. To simplify the problem still further, we neglect the effect of rotation and the gravitational pull of the earth also and consider a free flight of the rocket. We assume, therefore, that a rocket fired to move along the x-axis will continue to move along the x-axis itself.

Consider a rocket propelled by burning fuel. To write its equation of motion, we find the change in the momentum of the whole system in time interval Δt. Let M be the mass of the rocket and v its speed at time t. Then, in time interval Δt, the mass of the system is reduced by amount ΔM due to a burning of the fuel and expulsion of an equal amount of mass of gas. As a result of this reduction in mass, the velocity of the system increases by amount Δv. Let u be the velocity of the exhaust

gases relative to the rocket (Fig. 3.10). Then, the law of conservation of momentum gives

$$Mv = (M - \Delta M)(v + \Delta v) - \Delta M(u - v) \tag{3.152}$$

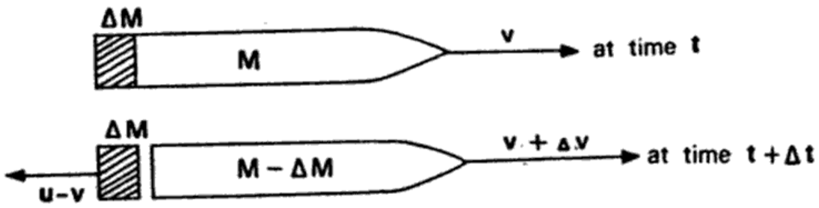


Fig. 3.10 Rocket motion

Simplifying equation (3.152) and retaining only the terms containing first-order infinitesimal quantities, we get

$$M\,\Delta v = u\,\Delta M$$

Dividing throughout by Δt and taking the limit as $\Delta t \to 0$, we get

$$M\dot{v} = -u\frac{dM}{dt} \tag{3.153}$$

where $\dot{v} = \frac{dv}{dt}$. The negative sign is added on the right-hand side to indicate that velocity v increases as mass M decreases.

Integrating equation (3.153) with respect to time, we get

$$\int_{v_0}^{v} dv = -u\int_{M_0}^{M_t}\frac{dM}{M}$$

or

$$v = v_0 - u\ln\frac{M_t}{M_0} \tag{3.154}$$

where v and M_t are the velocity and mass of the system at instant t and v_0 and M_0 those at $t = 0$.

Let us suppose that the fuel is burnt at constant rate $\frac{dM}{dt} = b$ and it lasts for time T. If the mass of the vehicle is M_v and that of the fuel at $t = 0$ is M_f, then

$$M_0 = M_v + M_f \tag{3.155}$$

The mass of the vehicle-fuel system at any instant t can be written as

$$M_t = M_v + M_f\left(1 - \frac{t}{T}\right) = M_0 - M_f\,\frac{t}{T}, \text{ for } 0 \leqslant t \leqslant T$$

and

$$M_t = M_v, \quad \text{for} \quad t \geqslant T$$

Substituting the value of M_t in equation (3.154), we get

$$v = \frac{dx}{dt} = v_0 - u\ln\left(1 - \frac{M_f}{M_0}\frac{t}{T}\right) \tag{3.156}$$

Integrating equation (3.156), with respect to time, we get

$$x = x_0 + v_0 t - u\int_0^t \ln\left(1 - \frac{M_f}{M_0}\frac{t}{T}\right)dt \tag{3.157}$$

The last integral of equation (3.157) can be evaluated by parts and we obtain

$$\int_0^t \ln\left(1 - \frac{M_f}{M_0}\frac{t}{T}\right) dt = \left(t - \frac{M_0 T}{M_f}\right) \ln\left(1 - \frac{M_f}{M_0}\frac{t}{T}\right) - t$$

Thus, the distance covered by the rocket in time t is given by

$$x = x_0 + v_0 t - u\left[\left(t - \frac{M_0 T}{M_f}\right) \ln\left(1 - \frac{M_f}{M_0}\frac{t}{T}\right) - t\right] \tag{3.158}$$

The rocket attains maximum velocity at $t = T$ when all its fuel is burnt out. This maximum velocity calculated from equation (3.156) is given by

$$\begin{aligned} v_{\max} &= v_0 - u \ln\left(1 - \frac{M_f}{M_0}\right) \\ &= v_0 + u \ln \frac{M_0}{M_v} \\ &= v_0 + u \ln\left(1 + \frac{M_f}{M_v}\right) \end{aligned} \tag{3.159}$$

From equation (3.159), it is clear that the larger the value of ratio $\frac{M_f}{M_v}$, the greater will be the maximum velocity attained by the rocket.

If the rocket is moving vertically upward and if the gravitational pull of the earth on it is assumed to be constant, then the equation of motion of the rocket can be written as

$$M\dot{v} = -u\frac{dM}{dt} - Mg$$

or

$$\dot{v} = -\frac{u}{M}\frac{dM}{dt} - g \tag{3.160}$$

Integrating equation (3.160) with respect to time, we get

$$v = -u \ln \frac{M_t}{M_0} - gt \tag{3.161}$$

Assuming the initial conditions as $x_0 = 0$ and $v_0 = 0$, we can derive an expression for the height attained by a rocket at time t and is given by

$$x = ut - \tfrac{1}{2}gt^2 - \left(t - \frac{M_0 T}{M_f}\right) \ln\left(1 - \frac{M_f}{M_0}\frac{t}{T}\right) \tag{3.162}$$

The rockets are normally expected to carry some load called the payload. Payload may be the load of a satellite which is to be placed in an orbit around the earth, or of bombs in the case of missiles. The payload and the body of the rocket have a fixed mass. Ratio $\frac{M_f}{M_v}$ has a practical limit and it cannot be increased indefinitely. A single rocket, i.e., one-stage rocket, therefore, will not attain high velocities that are required. Multistage rockets are used for this purpose.

QUESTIONS

1. What is a particle ? Can atom or earth be treated as a particle ? Explain.
2. In the equation of motion of system $\mathbf{F} = m\mathbf{a}$, what is represented by each side ? To answer this, consider a motion of a body falling towards the earth through the atmosphere.
3. When does Newton's third law break down ?
4. What is the force for which potential function $V(\mathbf{r})$ does not exist ?
5. How does Newton's second law govern the behaviour of (a) the linear momentum, and (b) the angular momentum, of a particle ? When are $\mathbf{p}$ and $\mathbf{L}$ conserved ?
6. In a projectile motion, when air resistance is negligible, is it necessary to consider three-dimensional motion instead of two-dimensional motion ?
7. Can the acceleration of a projectile be represented by a radial and a tangential component at each point of the motion ? If so, is it advantageous to represent it in this manner ?
8. A particle of mass m is moving with velocity $\mathbf{v}$. Under what conditions of m and $\mathbf{v}$ will the following apply ? (a) Classical mechanics, (b) quantum mechanics, and (c) relativistic mechanics.
9. Distinguish between centre of mass and centre of gravity.
10. Explain the idea of Newtonian relativity.
11. The determination of potential energy and kinetic energy is relative. Explain.
12. What is meant by an 'inertial mass' and 'gravitational mass' ? Is there any difference between the two ?
13. Newton's second law of motion is $\mathbf{F} = \dfrac{d}{dt}(m\mathbf{v})$. Under what condition can we write $\mathbf{F} = m\dfrac{d\mathbf{v}}{dt}$ and $\mathbf{F} = m\dfrac{d\mathbf{v}}{dt} + \dfrac{dm}{dt}\mathbf{v}$?
14. When is a force-field said to be conservative ? Give illustrations.
15. What is meant by terminal velocity ? Give illustrations.
16. You are given a system of N particles on which external force $\mathbf{F}$ is acting. Show that the centre of mass of the system behaves like a particle whose mass is equal to the total mass of the system and is acted upon by total external force $\mathbf{F}$.
17. Multistage rockets are used while launching satellites. Explain why
18. In the case of rocket motion, show that the greater the ratio M_f/M_e the greater is the maximum speed attained by the rocket.
19. Show that the centre of gravity coincides with the centre of mass when a body is in a uniform gravitational field. What will happen when force-field is non-uniform ?

20. Does frictional loss occur in the collision of one molecule with the other ? Explain.
21. Give some examples of motion that are approximately simple harmonic. Why are motions that are exactly simple harmonic, rare ?

PROBLEMS

1. A particle tied to an inextensible light string is rotating in a vertical plane. If its speed at the lowest point of the circle is v_0, find the minimum value of the speed v_m so that the particle will not leave the track. If v_0 is 0.775 m/s, find the point where the particle leaves the track.
2. A ball is tied to a string of length l and suspended from a nail in the wall. The ball is displaced from its equilibrium position and released when the string was horizontal. There is another nail vertically below the suspension and d is the distance between the nails. Show that $d \geqslant 0.6l$ if the ball completes a circle around the lower nail.
3. Consider a simple pendulum having massless inextensible rod of length l and having speed v_0 at displacement θ_0 less than $\frac{\pi}{2}$. Find the maximum values of v_0 at θ_0 (a) for $\theta = \frac{\pi}{2}$ to be reached by the pendulum, and (b) to keep the pendulum going in a vertical circle.
4. If the rod in problem 3 is elastic, show that it will be stretched at the lowest point by amount $\Delta l \simeq 3mg/k$, if $\Delta l \ll l$ and $\theta_0 = \frac{\pi}{2}$, $v_0 = 0$.
5. A particle of mass m falls along a frictionless track, the lower part of the track being circular. The particle starts from rest from point P which is at a vertical height of $5R$, where R is the radius of the circular part of the track. (a) What is the resultant force on it at point Q at a vertical height R and situated on the outer part of the circular part of the track? (b) At what height from the bottom should m be released if its force against the track at the top of the loop is equal to mg? Express the answers in terms of the speed at the bottom.
6 (a) Calculate the work done by force $\mathbf{F} = 4y\mathbf{i} - 2x\mathbf{j} - \mathbf{k}$ along helix $x = 4 \cos \theta$, $y = 4 \sin \theta$ and $x = 2\theta$ from $\theta = 0$ to $\theta = 2\pi$. (b) Calculate the work done by force $\mathbf{F} = 2x\mathbf{i} - 3z^2\mathbf{j} - y^2\mathbf{k}$ along the line $x = 2y = 4z$ from the origin to point (4, 2, 1).
7. A particle having initial velocity $v\mathbf{i}$ passes through a region in which there is electric field $E\mathbf{j}$ and magnetic field $B\mathbf{k}$. If the mass and the charge of the particle are m and e respectively, for what value of velocity $\dot{\mathbf{r}}$ will the particle move along the straight line?

8. A projectile shot from the ground has range R and the maximum height it reaches is H. Find the magnitude and the direction of its initial velocity (in the plane of its trajectory).
9. A particle has total energy E and the force on it is due to potential field $V(x)$. Show that the time taken by the particle to go from x_1 to x_2 is

$$t_2 - t_1 = \int_{x_1}^{x_2} \left[\frac{2E}{m} - V(x)\right]^{-1/2} dx$$

   if the motion is one-dimensional. Show that this is true only if the particle does not reverse its motion.
10. A billiard ball is dropped on a table with velocity v_0 and angular speed ω_0. At what time does slipping cease and rolling begin? Describe the subsequent motion of the ball.
11. A body is sliding down an inclined plane which is moving horizontally with constant velocity. Find the position of the body as a function of time.
12. A rotating sphere contracts slowly, due to internal forces, to $\frac{1}{n}$ of its original radius. What happens to its angular velocity? Show that increase in its energy is equal to work done during contraction.
13. Three particles of masses 2, 3 and 5 units move under the influence of a force-field so that their position vectors relative to a fixed coordinate system are given, respectively, by

$$\mathbf{r}_1 = 2t\mathbf{i} - 3\mathbf{j} - t^2\mathbf{k}$$
$$\mathbf{r}_2 = (t+1)\mathbf{i} - 3t\mathbf{j} - 4\mathbf{k}$$

   and $$\mathbf{r}_3 = t^2\mathbf{i} - t\mathbf{j} - (2t-1)\mathbf{k}$$

   where t is the time. Find (a) the total angular momentum of the system, and (b) the total external torque applied to the system with respect to the origin.
14. Three particles of mass 2, 1 and 3 units have the following position vectors:

$$\mathbf{r}_1 = 5t\mathbf{i} - 2t^2\mathbf{j} + (3t-2)\mathbf{k}$$
$$\mathbf{r}_2 = (2t-3)\mathbf{i} + (12-5t^2)\mathbf{j} + (4+6t+3t^2)\mathbf{k}$$

   and $$\mathbf{r}_3 = (2t-1)\mathbf{i} + (t^2+2)\mathbf{j} - t^3\mathbf{k}$$

   where t is time. Find (a) the velocity of the centre of mass of the system at $t = 1$, and (b) the total linear momentum of the system at $t = 1$.
15. A system of particles consists of particles of mass 3g located at point P (1, 0, −1), 5g at point Q (−2, 1, 3) and 2g at point R (3, −1, +1). Find the coordinates of the centre of mass of the system.
16. In a 2-stage rocket, the first stage gets detached after its fuel is used up. Each empty rocket (with neither fuel nor pay load) weighs 1/10th of the mass of the fuel it can contain. A pay load of 100 kg

is to be accelerated to a speed of 6000 m/s in a region free of external forces. The speed of exhaust gases is 1500 m/s. Find the choice of the masses of the two stages, including fuel so that the total mass at take-off is minimum. Also show that the required speed cannot be attained with a single stage rocket.

17. A rocket is initially at rest. It is driven by emitting photons. What fraction of the initial rest mass should be converted into energy if it is to reach speed v?

18. A raindrop of initial mass m_0 falls from rest through a cloud whose thickness is a. As the raindrop falls, it gains mass at rate b. The droplets of the cloud are at rest relative to the ground. The motion of the drop is resisted by a force proportional to its velocity. (a) Write down the differential equation of motion of the raindrop. (b) Find the velocity of the drop as it emerges from the cloud, if, during the passage, its mass has been doubled. (c) What will be the limiting velocity of the drop after it leaves the cloud, assuming that the air resistance outside the cloud is the same as that within?

19. Given force $\mathbf{F} = xy\mathbf{i} - y^2\mathbf{j}$, find the work done in moving a particle from (0, 0) to (2, 1).

20. Find the nature of the force-conservative or non-conservative if the work done is given by $W = x^2y - xz^3 - z$.

21. The equation of motion of a particle is

$$\frac{d^2\mathbf{r}}{dt^2} = \omega\mathbf{j} \times \frac{d\mathbf{r}}{dt}$$

where $\mathbf{j}$ is a constant unit vector and ω is a constant. Determine the motion.

22. A particle of mass m is moving under central force $f = \frac{k\mathbf{r}}{r^3}$, where $\mathbf{r}$ is the radius vector from the centre and k is a constant. Show that the vectors

$$\text{(i) } m\mathbf{r} \times \frac{d\mathbf{r}}{dt} \quad \text{and} \quad \text{(ii) } m\frac{d\mathbf{r}}{dt} \times \left(\mathbf{r} \times \frac{d\mathbf{r}}{dt}\right) + \frac{k\mathbf{r}}{r}$$

are constant and that they are orthogonal.

23. A bead is able to slide along a small wire in the form of a parabola. The parabola is rotating with constant angular velocity about its vertical symmetrical axis. Write down the equation of motion of the bead.

24. A particle of mass m moves under the influence of force $\mathbf{F}$ on the surface of a sphere of radius r. Write down the equation of motion of the particle.

25. A projectile is fired with velocity v_0 from a gun adjusted for a maximum range. It passes through two points P and Q whose heights above the horizontal are h each. Show that the separation of the

points is

$$x = \frac{v_0}{g}\sqrt{v_0^2 - 4gh}$$

26. A particle is projected vertically upward with velocity v_0 in a constant gravitational field. The medium offers a resistive force proportional to the square of the instantaneous velocity of the particle. If v_t denotes the terminal velocity, show that the velocity of the particle when it returns to the point from which it was projected is

$$v_0 v_t / \sqrt{v_0^2 + v_t^2}$$

27. Suppose that electrons could be added to earth and moon until repulsive force thus produced is just equal to balance the gravitational attraction. What would be the smallest total mass of electrons that would achieve this?

4

Inverse Square Law Field and Potential

There are four basic interactions in nature through which particles interact. They are termed strong, weak, electromagnetic and gravitational. Out of these, strong and weak interactions are very 'short-range' interactions. These are responsible for the binding of nucleus and the decay of nuclei and other particles respectively. The gravitational and electrostatic interactions are long-range interactions and are responsible for most of the phenomena we observe in nature. Thus, the motion of planets and satellites, and motion under the action of gravity are due to the gravitational field produced by some body. Electromagnetic interactions are responsible for friction, elasticity, surface tension, binding of electrons in an atom, etc.

Gravitational field is produced by the particles having mass. The heavier the mass, the stronger is the field produced. However, the gravitational field produced by a particle can only attract another particle having mass. Electrostatic field is produced by charges at rest. The charges are of two kinds—positive and negative. Two like charges—both positive or both negative—repel each other, whereas two unlike charges—one positive and the other negative—attract each other. According to the principle of superposition, the field produced by the positive charges can be reduced or nullified by the field produced by the negative charges. That is why an atom composed of an equal number of positively and negatively charged particles is neutral. But, we do not get masses of opposite nature—positive and negative. Hence, the gravitational field can only be increased by increasing the mass of the body which produces it.

Motion of charged particles produces a magnetic field. Some atoms and molecules have magnetic dipole moments which produce magnetic

domains in some materials. The alignment of these magnetic domains produces magnetism in magnets. The phenomenon of magnetism was explained by assuming the existence of positive or north poles and negative or south poles in a magnet. The two poles, however, cannot be separated and a free single pole called a monopole does not exist in nature. It is convenient to assume that the poles—like the charges—produce a magnetic field. Historically, the phenomenon of magnetism was studied with this assumption and the law of force between the two poles was established experimentally. The electric and magnetic fields are but special cases of the electromagnetic field produced by the moving charges.

The gravitational and electromagnetic fields are propagated with the velocity of light which is very large as compared to velocities of material bodies with which we are usually concerned. Hence, it can be assumed that the field is communicated with infinite velocity, i.e., the field acts instantaneously. A mass or a charge produces a gravitational and electrostatic field respectively which extends to infinity. The law of gravitational force was enunciated by Isaac Newton in his famous book '*Principia Mathematica*' published in 1687. In 1750, Michell showed that the law of force between two magnetic poles is an inverse square law. Coulomb, in 1765, used a torsion balance to demonstrate the inverse square law of force between two electric charges. All the three laws of force mentioned above are inverse square laws and hence have many common features. Since, magnetic monopoles do not exist, the magnetic field is produced only by a dipole or dipole distributions in materials. However, we can have a charge distribution like a mass distribution and hence, in this chapter, we shall discuss the common features of gravitational and electrostatic field. We shall also discuss properties of dipole and quadrupole which can be applicable to magnetic dipoles and quadrupoles.

4.1 LAWS OF GRAVITATIONAL AND ELECTROSTATIC FORCES

Newton's law of gravitation states that *every body in the universe attracts every other body with a force which is directly proportional to the product of their masses and inversely proportional to the square of their distance apart.*

If m_1 and m_2 are the masses of the two particles and r the distance between them, then the force of attraction F on m_2 due to m_1 which acts along the line joining the particles is

$$F \propto \frac{m_1 m_2}{r^2}$$

This can be transformed into a vector equation

$$\mathbf{F} = -G\frac{m_1 m_2}{r^2}\,\hat{\mathbf{e}}_r \tag{4.1a}$$

where G is the universal constant of gravitation and $\hat{\mathbf{e}}_r$ is a unit vector along the radius along which the force acts (Fig. 4.1). The negative sign is introduced to indicate the force of attraction. This also helps in

developing the theory of electrostatic and gravitational fields on parallel lines.

The magnitude of G has been determined by many scientists such as Lord Cavendish, Poynting and Boys. The presently-accepted value of G is

$G = (6.673 \pm 0.003) \times 10^{-11}$ Nm^2/kg^2

For a body near the surface of the earth, the force acting on it and directed towards the centre of the earth is given by

$$\mathbf{F} = m\mathbf{g}$$

and also by

$$\mathbf{F} = -\frac{GMm}{R^2}\,\hat{\mathbf{e}}_R$$

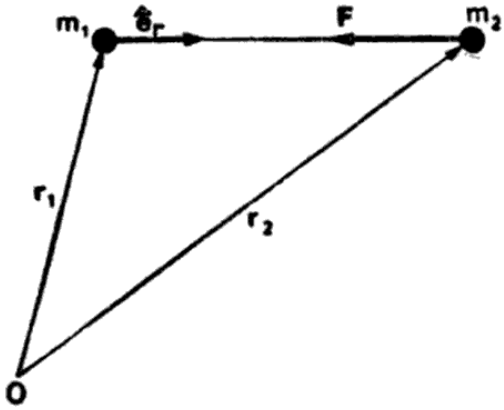


Fig. 4.1 Force of attraction between two masses: m_1 acts as a source

This gives

$$\mathbf{g} = -\frac{GM}{R^2}\,\hat{\mathbf{e}}_R \tag{4.2}$$

where M is the mass of the earth and R is its radius.

Equation (4.1a) can also be written in a vector form as follows:

Let $\mathbf{r}_1$ and $\mathbf{r}_2$ be the position vectors of the two particles with respect to some origin. Then, the gravitational force on the particle of mass m_2 due to that of mass m_1 is given by

$$\mathbf{F}_{1\to 2} = -\frac{Gm_1m_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3}(\mathbf{r}_1 - \mathbf{r}_2) \tag{4.1b}$$

Vector $(\mathbf{r}_1 - \mathbf{r}_2)$ gives the correct direction of the force and its magnitude is divided by extra factor $|\mathbf{r}_1 - \mathbf{r}_2|$ introduced in the denominator.

Equation (4.1b) can also be written as

$$\mathbf{F}_{1\to 2} = -\frac{Gm_1m_2}{|\mathbf{r}_1 - \mathbf{r}_2|^2}\,\hat{\mathbf{e}}_r$$

where $\hat{\mathbf{e}}_r$ is the unit vector in the direction of vector $(\mathbf{r}_1 - \mathbf{r}_2)$.

The law as formulated above is applicable only to particles or to bodies whose dimensions are negligible in comparison with the distance between them.

Force between two charges q_1 and q_2 separated by a distance is given by Coulomb's inverse square law which can be stated as

$$\mathbf{F} = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r^2}\,\hat{\mathbf{e}}_r \tag{4.3}$$

where charges q_1 and q_2 may be positive or negative and ϵ_0 is the permittivity of the vacuum. In equation (4.3), S.I. units are used for expressing all the quantities, viz. newton for force, coulomb for charge and meter for distance. Further,

$$\frac{1}{4\pi\epsilon_0} = 9 \times 10^9\,\frac{\text{Nm}^2}{\text{coul}^2} \quad \text{or} \quad \epsilon_0 = 8.9 \times 10^{-12}\,\frac{\text{coul}^2}{\text{Nm}^2}$$

In equation (4.3), when F is positive, it will represent repulsion and when F is negative, it will represent attraction between two charges.

It is interesting to compare the magnitudes of the gravitational and the Coulomb forces between two particles, say electrons. The ratio of the forces is

$$\frac{F_q}{F_m} = \frac{1}{4\pi\epsilon_0 G}\left(\frac{q}{m}\right)^2 \tag{4.4a}$$

For electron, $q = 1.6 \times 10^{-19}$C and $m = 9.1 \times 10^{-31}$ kg

Hence, ratio $\frac{F_q}{F_m}$ is

$$\frac{F_q}{F_m} \approx 4 \times 10^{42} \tag{4.4b}$$

Thus, the electrostatic forces are very much stronger than the gravitational forces. Hence, the gravitational interaction is neglected while considering the interaction of atomic or subatomic particles. However, the gravitational forces start becoming significant if the number of particles, i.e., atoms or molecules, is large. Another fact needs to be remembered in the case of charges and masses. The total mass and charge is conserved; a fact well established experimentally. It is only in high energy physics that conversion of mass into energy according to Einstein's mass-energy formula $E = mc^2$, has to be taken into account. Here E is the energy obtained by converting mass m into energy and c is the velocity of light. Under these conditions, the law of conservation of mass has to be changed into the law of conservation of mass and energy.

4.2 GRAVITATIONAL AND ELECTROSTATIC FIELDS AND POTENTIALS

Gravitational force $\mathbf{F}_m$ acting on a particle of mass m and situated at point $\mathbf{r}$ due to other particles having masses m_i and situated at points $\mathbf{r}_i$ is found out by the principle of superposition and is given by

$$\mathbf{F}_m = -\sum_i \frac{Gmm_i}{|\mathbf{r}_i - \mathbf{r}|^2}\,\hat{\mathbf{e}}_r \tag{4.5a}$$

If instead of point masses m_i, we have a continuous distribution of mass in space, we change Σm_i into $\iiint\rho(\mathbf{r}')\,d\tau'$. Then, equation (4.5a) assumes the form

$$\mathbf{F}_m = -\iiint \frac{Gm\rho(\mathbf{r}')\,d\tau'}{|\mathbf{r}' - \mathbf{r}|^2}\,\hat{\mathbf{e}}_{r'} \tag{4.5b}$$

where $\rho(\mathbf{r}')$ is the density and $d\tau'$ is a volume element.

Corresponding formulae for electrostatic force $\mathbf{F}_q$ on charge q due to a system of charges q_i are

$$\mathbf{F}_q = \sum_i \frac{\gamma qq_i}{|\mathbf{r}_i - \mathbf{r}|^2}\,\hat{\mathbf{e}}_r \tag{4.6a}$$

and

$$\mathbf{F}_q = \iiint \frac{\gamma q\rho(\mathbf{r}')\,d\tau'}{|\mathbf{r}' - \mathbf{r}|^2}\,\hat{\mathbf{e}}_{r'} \tag{4.6b}$$

where $\gamma = \dfrac{1}{4\pi\epsilon_0}$ and $\rho(\mathbf{r}')$ is the density of charge at point $\mathbf{r}'$.

Thus, we find that $\mathbf{F}_m$ (or $\mathbf{F}_q$) is proportional to mass m (or charge q).

We now define the *gravitational field intensity* or *gravitational field* $\mathbf{g}(\mathbf{r})$, *at any point* $\mathbf{r}$ *in space, due to any distribution of mass, as the gravitational force experienced by a unit mass situated at that point.* Thus

$$\mathbf{g}(\mathbf{r}) = \frac{\mathbf{F}_m}{m}$$

Hence, from equations (4.5a) and (4.5b), we can write

$$\mathbf{g}(\mathbf{r}) = -\sum_i \frac{Gm_i}{|\mathbf{r}_i - \mathbf{r}|^2}\,\hat{\mathbf{e}}_r \tag{4.7}$$

and

$$\mathbf{g}(\mathbf{r}) = -\iiint \frac{G\rho(\mathbf{r}')\,d\tau'}{|\mathbf{r}' - \mathbf{r}|^2}\,\hat{\mathbf{e}}_{r'} \tag{4.8}$$

Field $\mathbf{g}(\mathbf{r})$ has the dimensions of acceleration. It is the acceleration of the particle of mass m situated at point $\mathbf{r}$ on which only gravitational forces are acting.

To calculate the value of gravitational field $\mathbf{g}(\mathbf{r})$, we shall carry out the sum or the integration mentioned in equation (4.7) or (4.8). This is not a simple task as it involves the addition of a large number of vectors. We, therefore, define yet another quantity, viz. the *gravitational potential* at any point in a given gravitational field. The gravitational force between a pair of particles acts along the line joining the two particles and is conservative. In such a case, the potential energy for a pair of particles having masses m and m_i is defined by the formula

$$V_{mm_i} = -\frac{Gmm_i}{|\mathbf{r}_i - \mathbf{r}|} \tag{4.9}$$

Gravitational potential $\Phi(\mathbf{r})$ at any point $\mathbf{r}$ is defined as the potential energy, per unit mass, of a particle situated at point $\mathbf{r}$. Thus, we have

$$\Phi(\mathbf{r}) = \frac{V_m(\mathbf{r})}{m} \tag{4.10}$$

In the case of a system of particles

$$\Phi(\mathbf{r}) = -\sum_i \frac{Gm_i}{|\mathbf{r}_i - \mathbf{r}|} \tag{4.11}$$

Similarly, in the case of a continuous distribution of mass

$$\Phi(\mathbf{r}) = -\iiint \frac{G\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|}\,d\tau' \tag{4.12}$$

The negative sign in equations (4.9) and (4.11) is chosen arbitrarily to have zero potential energy at infinity. The particle always tends to move towards a point where its potential energy would be minimum.

The potential defined in this manner is a scalar point function. It is rather easy to calculate the potential than to calculate the gravitational field intensity. Knowing the potential function at a given point, the

gravitational field intensity at that point can be calculated by using the formula

$$\mathbf{g}(\mathbf{r}) = -\nabla\Phi(\mathbf{r}) \tag{4.13}$$

Conversely

$$\Phi(\mathbf{r}) - \Phi(\mathbf{r}_0) = -\int_{r_0}^{r} \mathbf{g}(\mathbf{r}')\cdot d\mathbf{r}' \tag{4.14}$$

where $\mathbf{r}_0$ represents the position vector of the initial configuration at which the potential is $\Phi(\mathbf{r}_0)$. Usually $|\mathbf{r}_0|$ is chosen to be infinity so that the corresponding point is far away from all the masses constituting the system and corresponding potential $\Phi(\mathbf{r}_0)$ can be taken to be zero. Thus, potential at infinity is zero (i.e. maximum), while at all finite distances, it is negative. With this choice, formula (4.14) becomes

$$\Phi(\mathbf{r}) = -\int_{\infty}^{r} \mathbf{g}(\mathbf{r}')\cdot d\mathbf{r}' \tag{4.15}$$

The above discussion is applicable to electrostatic fields also. Corresponding results can immediately be written down by noting the difference in symbols in equations (4.5) and (4.6).

Thus, the electric field intensity is given by

$$\mathbf{E}(\mathbf{r}) = \frac{\mathbf{F}_q}{q} \tag{4.6a}$$

$$= \sum \frac{\gamma q_i}{|\mathbf{r}_i - \mathbf{r}|^2}\hat{\mathbf{e}}_r \tag{4.7a}$$

$$= \iiint \frac{\gamma\rho(\mathbf{r}')\,d\tau'}{|\mathbf{r}' - \mathbf{r}|^2}\hat{\mathbf{e}}_{r'} \tag{4.8a}$$

Similarly, electrostatic potential is given by

$$\Phi(\mathbf{r}) = \frac{V_q(\mathbf{r})}{q} \tag{4.10a}$$

$$= \sum \frac{\gamma q_i}{|\mathbf{r}_i - \mathbf{r}|} \tag{4.11a}$$

$$= \iiint \frac{\gamma\rho(\mathbf{r}')\,d\tau'}{|\mathbf{r}' - \mathbf{r}|} \tag{4.12a}$$

The relation between $\mathbf{E}(\mathbf{r})$ and $\Phi(\mathbf{r})$ can be expressed as

$$\mathbf{E}(\mathbf{r}) = -\nabla\Phi(\mathbf{r}) \tag{4.13a}$$

$$\Phi(\mathbf{r}) = -\int_{\infty}^{r} \mathbf{E}(\mathbf{r}')\cdot d\mathbf{r}' \tag{4.15a}$$

4.3 LINES OF FORCE AND EQUIPOTENTIAL SURFACES

Consider a point charge (mass) around which there exists field $\mathbf{E}$ (or $\mathbf{g}$). Let us draw a line outward from the point charge (mass) such that the direction of the line at every point coincides with the direction of the field at that point. Such a line extends from the point charge (mass) to infinity and is called a line of force. A large number of such lines can be drawn extending from the point charge (mass) to infinity. Each line

gives the direction of **E** (or **g**) at the corresponding points. Thus, the lines of force are related to the direction of **E** (or **g**) at the point under consideration. We can introduce the *magnitude* of the field by defining the *density* of lines to be proportional to the magnitude of the field intensity in the region under consideration. Thus, the number of lines of force passing through a unit area drawn around the point perpendicular to the direction of intensity is, by convention, directly proportional to the magnitude of the field intensity at that point. The concept of lines of force is, therefore, a convenient way of visualizing the magnitude as well as the direction of the field intensity.

We have already defined the potential point function $\Phi(\mathbf{r})$ as the potential energy per unit charge (mass) at any point $\mathbf{r}$ in space. The equation

$$\Phi(\mathbf{r}) = \text{const} \tag{4.16}$$

defines a surface such that the potential at all points on it has the same magnitude. Such a surface is called an equipotential surface. Since, $\mathbf{E} = -\nabla\Phi(\mathbf{r})$ and $\Phi(\mathbf{r}) = \text{a}$ constant for an equipotential surface, $\mathbf{E}\cdot d\mathbf{r} = -d\Phi = 0$ for displacement $d\mathbf{r}$ along an equipotential surface. Hence, **E** will have no component along the equipotential surface. The lines of force will, therefore, be normal to the equipotential surface. If a charge (mass) moves on an equipotential surface, no work will be done by the field. Further, since the field intensity at a point is unique, the potential point function is single-valued and, therefore, the equipotential surfaces do not intersect each other. The equipotential surfaces surrounding a single isolated point charge (mass) are all spheres.

4.4 FIELDS AND POTENTIALS OF DIPOLE AND QUADRUPOLE

Two equal and opposite charges separated by a certain distance constitute an electric dipole. We have seen that the smallest magnetic structure that can exist independently is a magnetic dipole. But, we cannot get a gravitational dipole since negative mass producing opposite field does not exist in nature.

Consider two charges q and $-q$ separated by a small distance $\mathbf{a}$ (Fig. 4.2). The potential at any point $P(r, \theta, \varphi)$ due to this dipole is the sum of the potentials due to the two charges and is given by

$$\Phi(\mathbf{r}) = \frac{\gamma q}{|\mathbf{r} - \mathbf{a}|} - \frac{\gamma q}{|\mathbf{r}|} \tag{4.17}$$

From Fig. 4.2, $|\mathbf{r} - \mathbf{a}| = (r^2 - 2ra\cos\theta + a^2)^{1/2}$, where θ is the angle between $\mathbf{r}$ and $\mathbf{a}$. We assume that $r \gg a$. Hence

$$\begin{aligned}\frac{1}{|\mathbf{r} - \mathbf{a}|} &= \frac{1}{r}\left(1 - \frac{2a}{r}\cos\theta + \frac{a^2}{r^2}\right)^{-1/2} \\ &= \frac{1}{r}\left[1 - \tfrac{1}{2}\left(-\frac{2a}{r}\cos\theta + \frac{a^2}{r^2}\right) + \tfrac{3}{8}\left(-\frac{2a}{r}\cos\theta + \frac{a^2}{r^2}\right)^2 - \dots\right] \\ &= \frac{1}{r} + \frac{a}{r^2}\cos\theta + \frac{a^2}{r^2}\left(\tfrac{3}{2}\cos^2\theta - \tfrac{1}{2}\right)\dots \end{aligned} \tag{4.18}$$

In vector notation, we can write equation (4.18) as*

$$\frac{1}{|\mathbf{r}-\mathbf{a}|} = \frac{1}{r} + \frac{\mathbf{a}\cdot\mathbf{r}}{r^3} + \frac{3(\mathbf{a}\cdot\mathbf{r})^2 - a^2r^2}{2r^5} + \ldots \tag{4.19}$$

If we neglect terms containing a^2 or higher powers of a, we get, from equations (4.17) and (4.19)

$$\Phi(\mathbf{r}) = \gamma\frac{\mathbf{d}\cdot\mathbf{r}}{r^3} = \frac{\gamma d}{r^3}\cos\theta \tag{4.20}$$

where $\mathbf{d} = q\mathbf{a}$ is called the electric dipole moment.

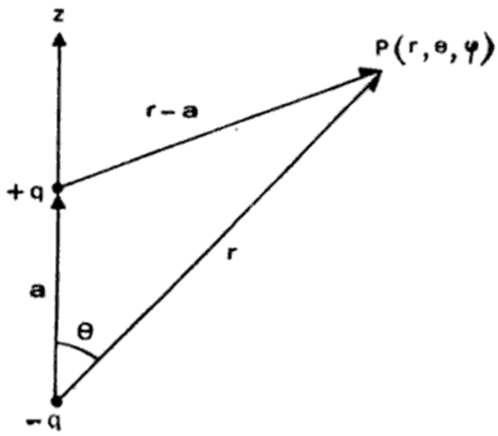


Fig. 4.2 An electric dipole

Thus, the potential due to a dipole falls off as the square of distance r. It should be remembered that the potential due to a point charge falls off as $\frac{1}{r}$. Using equation (4.13a) and the spherical polar form of ∇, we get the components of intensity at point $P(r, \theta, \varphi)$, Fig. (4.1), as

$$\left.\begin{aligned} E_r &= -\frac{\partial\Phi}{\partial r} = \frac{2\gamma d\cos\theta}{r^3} \\ E_\theta &= -\frac{1}{r}\frac{\partial\Phi}{\partial\theta} = \frac{\gamma d\sin\theta}{r^3} \\ \text{and}\qquad E_\varphi &= -\frac{1}{r\sin\theta}\frac{\partial\Phi}{\partial\varphi} = 0 \end{aligned}\right\} \tag{4.21}$$

Thus, the field intensity of a dipole varies inversely as the cube of the distance. It has no component along increasing φ direction.

The equation of the lines of force can be found by noting that intensity $\mathbf{E}$ and $d\mathbf{r}$, a displacement along the line of force at point P are parallel. Displacement $d\mathbf{r}$ along the line of force, if resolved along the radial and transverse directions, will have components dr and $rd\theta$

---

*The result of equation (4.19) is a well-known mathematical relation and can be written as

$$\frac{1}{|\mathbf{r}-\mathbf{a}|} = \sum_{l=0}^{\infty}\frac{a^l P_l(\cos\theta)}{r^{l+1}}$$

where $P_l(\cos\theta)$ are the Legendré polynomials of degree l.

(Fig. 4.3a). To find the equation of lines of force, let us take radial component E_r and transverse component E_θ to be proportional to dr and $r d\theta$ respectively. Then, we get

$$\frac{E_\theta}{E_r} = r\frac{d\theta}{dr}$$

But, by equation (4.21), we have

$$\frac{E_\theta}{E_r} = \frac{\sin\theta}{2\cos\theta}$$

Comparing these expressions, we get

$$\frac{dr}{r} = \frac{2\cos\theta\, d\theta}{\sin\theta} = \frac{2d(\sin\theta)}{\sin\theta}$$

Integration of this yields

$$r = A\sin^2\theta \tag{4.22}$$

which is the equation of lines of force. Here constant A is a parameter which varies from one line of force to another.

The field and potential due to a dipole are shown in Fig. 4.3b in which

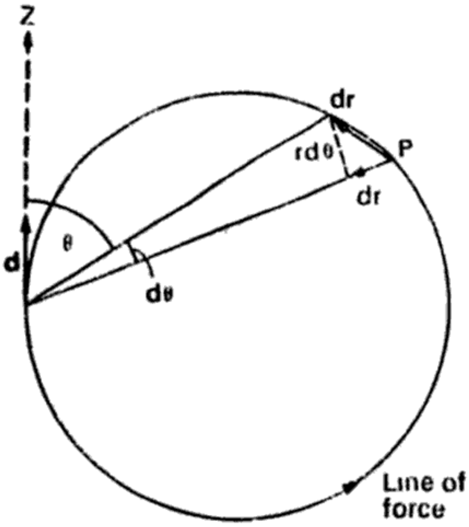


Fig. 4.3a Lines of force at the point $P(r, \theta)$

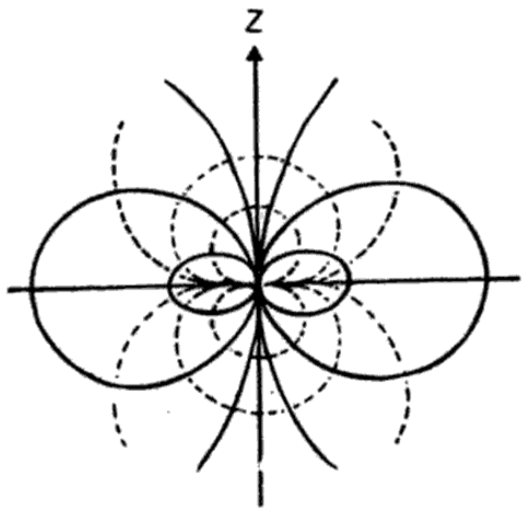


Fig. 4.3b Field and potential of an electric dipole. Solid lines—lines of force; dotted lines—equipotential lines

the solid lines represent intensity $\mathbf{E}$ and the dotted lines represent equipotential surfaces Φ = constant. The equipotential surfaces are obtained by rotating the equipotential lines about the z-axis.

Consider now two dipoles of equal dipole moments—one placed at the origin and the other at a distance from it. Let dipole moments $d = qa$ be oppositely directed as shown in Fig. 4.4. This forms a quadrupole. Using the results obtained for a dipole, we can get the potential of the quadrupole field. Thus,

$$\Phi(\mathbf{r}) = \frac{\gamma\mathbf{d}\cdot(\mathbf{r}-\mathbf{a})}{|\mathbf{r}-\mathbf{a}|^3} - \frac{\gamma\mathbf{d}\cdot\mathbf{r}}{r^3} \tag{4.23}$$

By neglecting the terms of order $\left(\frac{a}{r}\right)^2$ and higher order, we get

$$\frac{1}{|\mathbf{r}-\mathbf{a}|^3} = \frac{1}{r^3}\left(1 - \frac{2\mathbf{a}\cdot\mathbf{r}}{r^2} + \frac{a^2}{r^2}\right)^{-3/2} \approx \frac{1}{r^3} + \frac{3\mathbf{a}\cdot\mathbf{r}}{r^5} \tag{4.24}$$

Hence, to this approximation, quadrupole potential is

$$\Phi(\mathbf{r}) = \gamma\,\frac{3(\mathbf{d}\cdot\mathbf{r})(\mathbf{a}\cdot\mathbf{r}) - (\mathbf{d}\cdot\mathbf{a})r^2}{r^5} \tag{4.25}$$

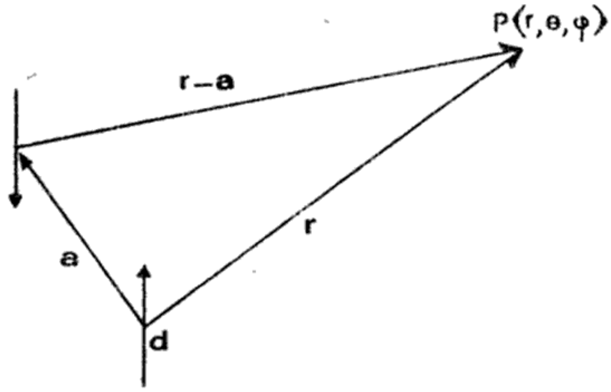


Fig. 4.4a Electric dipoles oppositely oriented

In the special case, when both dipoles are along the common direction, say the z-axis, as shown in Fig. 4.4b, and placed in an 'end-on' position such that

$$\mathbf{d}\cdot\mathbf{r} = dr\cos\theta,\ \mathbf{a}\cdot\mathbf{r} = ar\cos\theta,\ d = qa \text{ and } \mathbf{d}\cdot\mathbf{a} = da$$

We have

$$\Phi(\mathbf{r}) = \frac{\gamma Q}{4r^3}(3\cos^2\theta - 1) \tag{4.26}$$

where $Q = 4da$ is called the electric quadrupole moment. Factor 4 is added to simplify some later formulae. Thus, the quadrupole potential

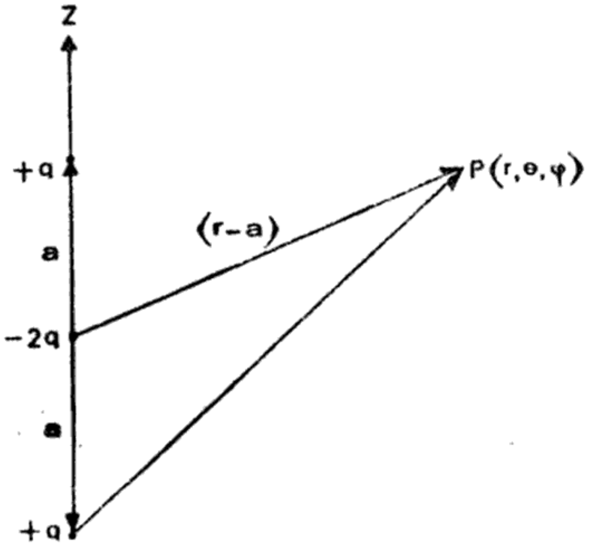


Fig. 4.4b Two electric dipoles along the common direction

varies inversely as the cube of the distance. The components of the field intensity are

$$\left.\begin{aligned} E_r &= \frac{3Q}{4r^4}(3\cos^2\theta - 1) \\ E_\theta &= \frac{3Q}{2r^4}\cos\theta\sin\theta \\ E_\varphi &= 0 \end{aligned}\right\} \tag{4.27}$$

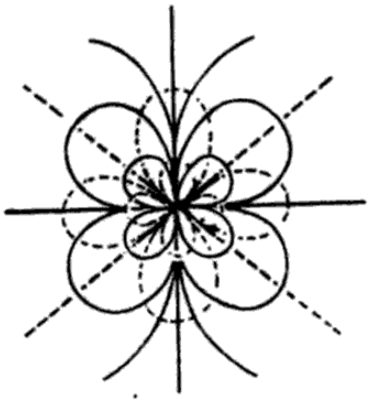


Fig. 4.5 Field and potential of a quadrupole

Thus, the field intensity is found to vary inversely as the fourth power of the distance. The variation of the quadrupole field and potential with distance is shown in Fig. 4.5.

The concepts of dipole and quadrupole can be extended to an assembly of a large number of equal positive and negative charges located in a small region. Such an arrangement is known as multipole and is helpful while considering the potential and field due to such a charge distribution.

4.5 POTENTIAL AND INTENSITY DUE TO A CHARGE DISTRIBUTION IN THE FORM OF A SPHERICAL SHELL

(a) Thin Shell

Consider a distribution of charge q (or mass m) in the form of a thin spherical shell of radius a and thickness da. We wish to find out the potential and intensity at any point P at distance r from the centre of the spherical charge distribution. The potential is given by equation (4.12a) wherein $\rho(r')$ is the constant charge density. The total charge will then be

$$dq = 4\pi a^2\, da\, \rho \tag{4.28}$$

It is convenient to choose the centre of the sphere as the origin and use spherical polar coordinates in which volume element $d\tau'$ is

$$d\tau' = r'^2\, dr' \sin\theta'\, d\theta'\, d\varphi'$$

Hence, the volume element of the shell of radius 'a' between coordinates θ' and $\theta' + d\theta'$ and φ' and $\varphi' + d\varphi'$ will be

$$d\tau' = a^2\, da \sin\theta'\, d\theta'\, d\varphi'$$

Let us choose the straight line joining the centre of the sphere to point ρ as the z-axis. Then, the potential is

$$d\Phi(\mathbf{r}) = \gamma\rho a^2\, da \int_0^\pi\int_0^{2\pi} \frac{\sin\theta'\, d\theta'\, d\varphi'}{(r^2 - 2ar\cos\theta' + a^2)^{1/2}}$$

Integrating over φ' gives factor 2π. Integration with respect to θ' can be performed by substituting

$$w = \cos\theta' \quad \text{and} \quad dw = -\sin\theta'\, d\theta'$$

so that $-1 < w < 1$. Thus

$$d\Phi(\mathbf{r}) = 2\pi\gamma\rho a^2\, da \int_{-1}^{+1} \frac{dw}{(r^2 - 2arw + a^2)^{1/2}}$$

where the negative sign is absorbed in the limits. Integration yields

$$d\Phi(\mathbf{r}) = 2\pi\gamma\rho a^2\, da \left[\frac{|r+a| - |r-a|}{ar}\right]$$

$$= \frac{\gamma\, dq}{2ar}\,[|r+a| - |r-a|] \tag{4.29}$$

When point P is outside the shell, i.e., when $r \geqslant a$, we get

$$d\Phi(\mathbf{r}) = \frac{\gamma\, dq}{r} \tag{4.30}$$

The potential at points outside the shell varies inversely as distance r. When point P is inside the shell, i.e., when $r < a$, we get

$$d\Phi(\mathbf{r}) = \frac{\gamma\, dq}{a} \tag{4.31}$$

Thus, the potential at points inside the shell is constant, and at points outside the shell the potential is the same as that for point charge dq situated at the centre of the shell. In other words, the intensity of the field at the points outside the shell is the same as if all the charge on the shell is concentrated at its centre.

The intensity of the field can now be computed by using equation (4.13a). This gives

$$\left.\begin{aligned} \mathbf{E}(\mathbf{r}) &= -\frac{\gamma\, dq}{r^2}\,\hat{\mathbf{e}}_r, \quad \text{for } r \geqslant a \\ &= 0 \qquad\qquad\quad \text{for } r < a \end{aligned}\right\} \tag{4.32}$$

(b) Thick Shell

We shall now extend the results obtained for a thin shell to a charge distribution in the form of a thick shell. Let r_1 and r_2 be the internal and external radius of the shell respectively.

We imagine the thick shell to be divided into a large number of thin concentric shells, each one contributing potential $d\Phi(\mathbf{r})$ to total potential $\Phi(\mathbf{r})$ at point P outside the shell. Thus

$$\Phi(\mathbf{r}) = \int d\Phi(\mathbf{r})$$

Since, $d\Phi(\mathbf{r}) = \dfrac{\gamma\, dq}{r} = \dfrac{4\pi\gamma r'^2\, dr'\, \rho}{r}$, where r' and dr' are the radius and the thickness of one of the elementary thin shells into which the thick shell is divided. Note that distance r of point P from the centre of the shell is constant. Thus

$$\Phi(\mathbf{r}) = \frac{4\pi\gamma\rho}{r} \int_{r_1}^{r_2} r'^2\, dr'$$

$$= \tfrac{4}{3}\pi \frac{\gamma\rho}{r} (r_2^3 - r_1^3)$$

$$= \frac{\gamma q}{r} \tag{4.33}$$

where the total charge q is

$$q = \tfrac{4}{3}\pi\rho(r_2^3 - r_1^3) \tag{4.34}$$

Thus, as far as the potential at a point on the surface of the shell or outside the shell is concerned, the charge distribution behaves as if all the charge is concentrated at the centre of the shell.

If point P is inside the shell and at distance r from the centre of the shell such that $r < r_1$, the total potential at P is obtained by summing constant potentials $d\Phi = \gamma\, dq/r$ due to a thin shell of radius r' into which the thick shell has been subdivided. Thus, the potential inside the cavity (i.e. when $r < r_1$) is

$$\Phi(\mathbf{r}) = 4\pi\gamma\rho \int_{r_1}^{r_2} r'\, dr'$$

$$= 2\pi\gamma\rho(r_2^2 - r_1^2)$$

$$= \tfrac{3}{2}\gamma q \frac{(r_2 + r_1)}{(r_2^2 + r_1 r_2 + r_1^2)} \tag{4.35}$$

where we have used equation (4.34) for the total charge. Equation (4.35) shows that the potential inside the shell is constant. By putting $r = r_1 = r_2$ in equation (4.35), we get the result of equation (4.31).

If point P is situated within the shell, then the potential at P is the sum of (i) potential due to a thick spherical shell of internal and external radii r_1 and r on the outer surface of which is situated point P and (ii) that due to a thick spherical shell of internal and external radii r and r_2 on the inner surface of which is situated point P.

Hence, the potential at P is

$$\Phi(\mathbf{r}) = \frac{4\pi\gamma\rho}{r} \int_{r_1}^{r} r'^2\, dr' + 4\pi\gamma\rho \int_{r}^{r_2} r'\, dr'$$

$$= \frac{4\pi\gamma\rho}{3r} (r^3 - r_1^3) + 2\pi\gamma\rho(r_2^2 - r^2)$$

$$= 4\pi\gamma\rho \left[\frac{r_2^2}{2} - \frac{r_1^3}{3r} - \frac{r^2}{6}\right] \tag{4.36}$$

When $r \to r_1$, we get the result of equation (4.35). Thus, the potential is a continuous function.

The intensity of the field is given by

$$\left.\begin{aligned} \mathbf{E}(r > r_2) &= -\frac{\partial\Phi(\mathbf{r})}{\partial r} = \frac{\gamma q}{r^2}\,\hat{\mathbf{e}}_r \\ \mathbf{E}(r < r_1) &= 0 \\ \mathbf{E}(r_1 < r < r_2) &= -\frac{4\pi\gamma\rho}{3}\left(\frac{r_1^3}{r^2} - r\right)\hat{\mathbf{e}}_r \end{aligned}\right\} \tag{4.37}$$

The intensity of the field is also a continuous function and obeys the inverse square law outside the shell.

The variation of potential $\Phi(\mathbf{r})$ and intensity $\mathbf{E}(\mathbf{r})$ is represented in Fig. 4.6.

We can obtain results for the uniform charge distribution in a sphere by putting $r_1 = 0$. Thus, from equation (4.37), we find that the intensity of the field decreases linearly with the distance.

The above discussion is also applicable to a gravitational field. In that case it is only necessary to replace γ by $-G$ and q by m.

As far as the effect of the uniform spherical distribution of charge or mass at points outside the sphere is concerned, the charge or mass can be replaced by a point charge or mass to be placed at the centre of the sphere. That is why we can take the sun as a point while considering the motion of a planet around the sun.

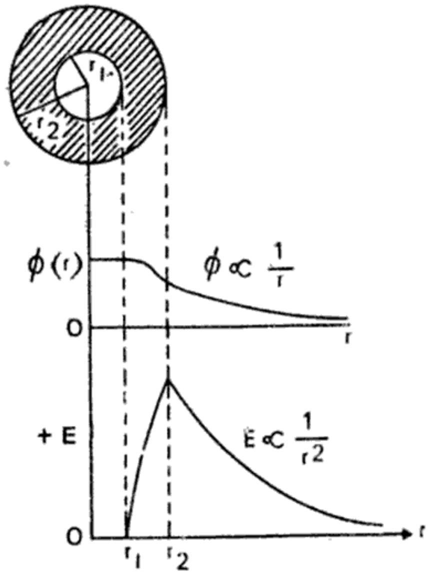


Fig. 4.6 Variation of potential and field due to a charge distribution in the form of a thick spherical shell

4.6 POTENTIAL DUE TO A CHARGE DISTRIBUTION AT LARGE DISTANCES

We can calculate the potential exactly by knowing the distribution of charge or mass in very few cases. But, if the distribution is of an arbitrary nature, we cannot evaluate the potential exactly and we have to employ the 'approximation method'. Let us consider an arbitrary charge distribution whose potential is given by equations (4.11) and (4.12). Let us also assume that $|\mathbf{r}'|$ is small when compared to $|\mathbf{r}|$, where r is the distance at which the potential is to be evaluated.

From equation (4.12a), we can write

$$\Phi(\mathbf{r}) = \int \frac{\gamma \rho(\mathbf{r}')\, d\tau'}{|\mathbf{r} - \mathbf{r}'|}$$

and as $r \gg r'$, by using equation (4.19), we get

$$\Phi(\mathbf{r}) = \frac{\gamma}{r} \int \rho(\mathbf{r}')\, d\tau' + \frac{\gamma}{r^2} \int \rho(\mathbf{r}')(\mathbf{r} \cdot \mathbf{r}')\, d\tau'$$

$$+ \frac{\gamma}{2r^5} \int \rho(\mathbf{r}')\{3(\mathbf{r} \cdot \mathbf{r}')^2 - r^2 r'^2\}\, d\tau' + \ldots \quad (4.38)$$

$$= \Phi_0(\mathbf{r}) + \Phi_1(\mathbf{r}) + \Phi_2(\mathbf{r}) + \ldots \quad (4.39)$$

The leading term in the expansion is

$$\Phi_0(\mathbf{r}) = \frac{\gamma q}{r} \quad (4.40)$$

where

$$q = \int \rho(\mathbf{r}') \, d\tau' = \sum q_i \tag{4.41}$$

represents the total charge. The term $\Phi_0(\mathbf{r})$ gives the potential at r due to the total charge q placed at the origin. Thus, for a large distance and to the lowest approximation, the charge distribution behaves as if charge is concentrated at a point. This term is called a 'monopole'. We have seen earlier that if the charge distribution is spherically symmetric, the total charge or mass behaves as if it is concentrated at the centre of the sphere as far as the evaluation of the potential and intensity at outside points is concerned. Thus, the successive terms in the series expansion of equation (4.39) will be the measures of deviation from spherical symmetry.

The second term in the series [equation (4.39)] is

$$\Phi_1(\mathbf{r}) = \gamma \frac{\mathbf{d} \cdot \mathbf{r}}{r^3} \tag{4.42}$$

where

$$\mathbf{d} = \int \rho(\mathbf{r}')\mathbf{r}' \, d\tau' \quad \text{or} \quad \mathbf{d} = \sum q_j \mathbf{r}_j \tag{4.43}$$

represents the 'dipole moment' of the charge distribution. Thus, the second term is the 'dipole term'.

Let $\mathbf{R}$ be the position vector of the *centre of charge* of the charge distribution when the total charge is non-zero. We shall define the centre of charge on the same lines as we defined the centre of mass. Here, however, we have to account for the existence of positive and negative charges. Let the position vector of charge q_j be

$$\mathbf{r}_j = \mathbf{R} + \mathbf{r}'_j$$

where $\mathbf{r}'_j$ is the position vector of q_j relative to the centre of charge. The dipole moment then becomes

$$\mathbf{d} = \sum q_j \mathbf{r}_j = \sum q_i \mathbf{R} + \sum q_j \mathbf{r}'_j \tag{4.44}$$

or

$$\mathbf{d} = q\mathbf{R} + \mathbf{d}' \tag{4.45}$$

where

$$\mathbf{d}' = \sum q_j \mathbf{r}'_j \tag{4.46}$$

Thus, the dipole moment will change if we shift the origin. From equation (4.44), we get

$$\begin{aligned} \mathbf{R} &= [\sum q_j \mathbf{r}_j - \sum q_j \mathbf{r}'_j] / \sum q_j \\ &= \frac{\mathbf{d} - \mathbf{d}'}{q} \end{aligned} \tag{4.47}$$

and can be made zero by choosing the centre of charge as the origin (i.e. $\mathbf{R} = 0$). If, however, the total charge is zero, the above result becomes indeterminate. However, equation (4.46) always determines the dipole moment unambiguously.

y-values) of the distribution. Thus, the value of Q will be a measure of deviation from the spherically symmetric distribution.

As an example, let us evaluate the quadrupole moment of the earth which is approximately an oblate spheroid. The equatorial radius a of the earth exceeds its polar radius b by about 25 km. We can define the oblateness as

$$\epsilon = \frac{a-b}{a} \simeq \frac{1}{300} \tag{4.53}$$

If we assume density ρ of the earth to be constant, we get

$$Q = 2\rho \int z'^2\, dx'\, dy'\, dz' - \rho \int (x'^2 + y'^2)\, dx'\, dy'\, dz'$$
$$= 2\rho\pi a^2 \int_{-b}^{b} z'^2\left(1 - \frac{z'^2}{b^2}\right) dz' - 2\pi\rho \int_{-b}^{b} \frac{x'^4}{4}\, dz'$$
$$= \frac{2M}{5}(b^2 - a^2)$$

where we have used $\rho = \dfrac{3M}{4\pi a^2 b}$ and $\dfrac{x'^2}{a^2} + \dfrac{y'^2}{b^2} = 1$. Hence, using equation (4.53), we get the quadrupole moment of earth as

$$Q = -\tfrac{2}{5}M(a^2 - b^2) \simeq -\tfrac{4}{5}Ma^2\epsilon \tag{4.54}$$

In fact, the density of the earth is larger at the centre and hence smaller values of r' will contribute greater amounts towards Q than the larger values of r'. We should, therefore, expect a smaller value of Q than that given by equation (4.54). The gravitational potential of the earth is

$$\Phi(\mathbf{r}) = -\frac{GM}{r} + \frac{GMa^2\epsilon}{5r^3}(3\cos^2\theta - 1) \tag{4.55}$$

In order to find the potential at any external point due to a charge (or mass) distribution, successively better approximation is obtained by finding potentials due to monopole, dipole moment, quadrupole moment, etc. of the distribution.

4.7 FIELD EQUATIONS

We have seen that by knowing the charge (or mass) distribution, we can evaluate the potential and the intensity of the field, by using equations (4.12) and (4.13). This process involves integration of potentials and fields due to constituent charges (or masses). When the charge (or mass) distribution is not known, it is convenient to solve differential equations by knowing the boundary conditions. We shall obtain equivalent differential equations by using Gauss' theorem.

Consider charge q at point O and let a closed surface be drawn around it (Fig. 4.7). Consider small element $d\sigma$ of the area on the surface situated at distance r from O. Then the intensity of field at a point on $d\sigma$ is given by

$$\mathbf{E} = +\frac{\gamma q}{r^2}\,\hat{\mathbf{e}}_r \tag{4.56}$$

where $\hat{\mathbf{e}}_r$ is along $\mathbf{r}$ from O. The component of $\mathbf{E}$ at right angles to the surface of $d\sigma$ is

$$E_n = +\frac{\gamma q}{r^2}\cos\theta \tag{4.57}$$

where θ is the angle between $\mathbf{E}$ and $d\boldsymbol{\sigma}$.

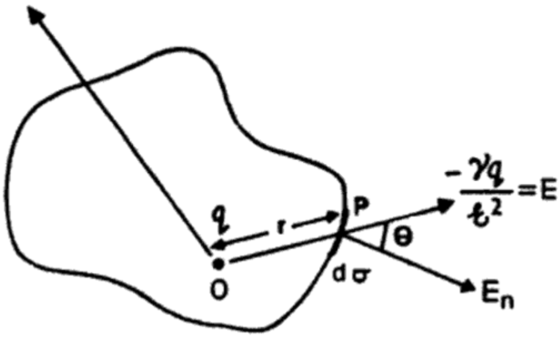


Fig. 4.7 Gauss' law in **electrostatics**

The flux through $d\sigma$ is, therefore, given by

$$E_n\, d\sigma = +\frac{\gamma q}{r^2}\cos\theta\, d\sigma \tag{4.58}$$

But, $d\sigma \cos\theta/r^2$ is solid angle $d\Omega$ subtended by area $d\sigma$ at O. Thus

$$E_n\, d\sigma = +\gamma q\, d\Omega \tag{4.59}$$

Integrating equation (4.59) over the given closed surface, we get

$$\int_\sigma E_n\, d\sigma = +\gamma q \int_\sigma d\Omega = 4\pi\gamma q \tag{4.60}$$

If the given surface has some portions such that a line drawn from O intersects it more than once, this number must be an odd number. Then, the contributions of such surfaces occurs once only in the integration, the even pairs of areas cancelling out their contributions.

If a large number of charges is present within the surface, we have, the total outward flux

$$\begin{aligned}\int E_n\, d\sigma &= \sum_i +\gamma q_i \int d\Omega\\ &= 4\pi\gamma q\end{aligned} \tag{4.61}$$

where $q = \sum q_i$.

For a continuous distribution of charge within the closed surface, we write the above equation as

$$\int_\sigma \mathbf{E}\cdot d\boldsymbol{\sigma} = \int_\tau 4\pi\gamma\rho(\mathbf{r})\, d\tau \tag{4.62}$$

where $d\tau$ is a small volume element and $\rho(\mathbf{r})$ is the charge density of the distribution. This is the well-known Gauss' law, according to which *the total outward flux through any closed surface is $4\pi\gamma$ times the total charge inside the surface.*

To illustrate the use of Gauss' law, consider a homogeneous sphere of

mass m. Let us find the force on a unit mass placed at distance r from the centre of the sphere and situated outside the sphere (Fig. 4.8). On account of the symmetry of the problem, we choose a spherical surface of radius r concentric with the sphere.

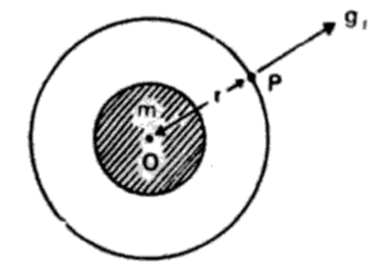


Fig. 4.8 Gravitational field due to a homogeneous sphere

Now, by Gauss' law

$$4\pi r^2 g_n = -4\pi Gm$$

or
$$g_n = -\frac{Gm}{r^2} \tag{4.63}$$

where we have replaced E_n by g_n, γ by $-G$ and q by m to suit the gravitational field. Since, the problem has a spherical symmetry, g_n is in the direction of vector $\mathbf{g}$ itself.

We now wish to find out a differential equation satisfied by field $\mathbf{E}(\mathbf{r})$ and potential $\Phi(\mathbf{r})$. Since

$$\mathbf{E} = -\nabla\Phi$$

it follows that

$$\nabla \times \mathbf{E} = 0 \tag{4.64}$$

This equation gives us three component equations in the rectangular cartesian coordinates, viz.

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = 0, \quad \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = 0, \quad \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = 0 \tag{4.65}$$

These equations alone do not determine the field uniquely. This is because these equations are satisfied by any conservative field. Hence, to determine the particular field, we need one more equation that relates field $\mathbf{E}$ to the distribution of charge that determines it. This relation is supplied by Gauss' law given in equation (4.62) above.

Converting the left-hand side of equation (4.62) into a volume integral by Gauss' theorem, we get

$$\int_\tau \nabla\cdot\mathbf{E}\, d\tau = +\int_\tau 4\pi\gamma\rho\, d\tau \tag{4.66}$$

or
$$\int_\tau (\nabla\cdot\mathbf{E} - 4\pi\gamma\rho)\, d\tau = 0 \tag{4.67}$$

Equation (4.64) must hold for any volume τ. This will be true only if

$$\nabla\cdot\mathbf{E}(\mathbf{r}) = +4\pi\gamma\rho(\mathbf{r}) \tag{4.68}$$

or
$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = 4\pi\gamma\rho(x, y, z) \tag{4.69}$$

in the component form.

If $\rho(x, y, z)$ is known, equation (4.69) together with equation (4.65) determines the field uniquely in the region of interest. However, the boundary conditions at the surface must be known. Substituting

$\mathbf{E} = -\nabla\Phi$ in equation (4.68), we get

$$\nabla^2\Phi(\mathbf{r}) = -4\pi\gamma\rho(\mathbf{r}) \tag{4.70}$$

i.e.

$$\frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = -4\pi\gamma\rho(x, y, z) \tag{4.71}$$

This equation alone determines potential $\Phi(x, y, z)$ if the boundary conditions are known. Equations (4.70) and (4.71) is called Poisson's equation. If $\rho = 0$ in the region of interest, we get

$$\nabla^2\Phi = 0 \tag{4.72}$$

This is called Laplace's equation.

Knowing charge distribution $\rho(\mathbf{r})$ in the region of interest and the boundary conditions at the surface instead of the charge distribution outside, we can find out potential inside the region by solving equations (4.70) and (4.72). While evaluating the potential by equations (4.12) and (4.13) we are required to know the charge distribution in all the space.

QUESTIONS

1. Consider a hollow spherical shell. How does the gravitational potential inside compare with that on the surface? What is the gravitational field-strength inside?
2. Can we regard gravitational force as a fictitious force arising from the acceleration of our reference frame relative to an inertial reference frame, rather than a real force?
3. If magnetic monopoles existed, would $\nabla \cdot \mathbf{B}$ be equal to zero?
4. What would $\oint \mathbf{H} \cdot d\mathbf{r}$ represent if magnetic monopoles existed?
5. Why is the gravitational potential always negative?
6. Explain the concept of lines of force. How is it useful in visualising the magnitude and direction of field intensity?
7. Equipotential surfaces do not intersect each other. Comment.
8. What is meant by a quadrupole? Define the quadrupole moment.
9. When does a distribution of equivalent dipoles produce a finite charge density?
10. An electric dipole is placed in a non-uniform electric field. Is there a net force on it?
11. An electric dipole has its dipole moment $\mathbf{p}$ aligned with uniform external field $\mathbf{E}$. Is the equilibrium stable or unstable?
12. Explain why a spherically symmetric distribution of mass yields a radially directed field intensity.
13. Show that the quadrupole moment of earth is nearly $-\frac{4}{5}Ma^2\epsilon$, where a is the equatorial radius of the earth.

14. Using Gauss' Law, find the intensity of gravitational field due to a homogeneous sphere of mass M.

PROBLEMS

1. Two similar balls are suspended from a common point by means of strings of equal lengths. If the two balls carry equal like charges, obtain an expression for the angle subtended by one of the strings with the vertical.
2. Obtain the charge distribution which yields the field
$$\mathbf{E} = A(2z\mathbf{k} - x\mathbf{i} - y\mathbf{j})$$
where A is a constant. Also find the potential.
3. A charge of 4 units is placed at point P and -1 unit at point Q. Within what angle of line PQ do the lines of force leave point P if these lines are to end on Q?
4. Two similar charges Q each are placed at the two extremities of straight line AB. A third charge Q is placed at the mid-point of AB and is allowed to oscillate along AB. Find, to the first approximation, the restoring force acting on this charge.
5. Find the charge on an area of one square kilometer of the earth's surface if an electric field of 10^4 V/m is directed vertically downward near the earth's surface?
6. Find the electric field at a point just outside a cylindrically symmetric charge distribution.
7. Find the charge that produces field
$$\mathbf{E} = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$$
Also find the potential that describes this field.
8. What charge distribution would produce the Yukawa potential
$$\Phi = \frac{q}{4\pi\epsilon_0}\frac{e^{-r/a}}{r}?$$
9. Show that the equipotential surfaces of a thin rod of finite length and carrying a charge are ellipsoids of revolution.
10. If the intensity of the field is independent of the radial distance within the sphere, find the function which describes density $\rho = \rho(\mathbf{r})$ of the sphere.
11. A particle moves under the action of force $F = -k^2/x^3$. Show that the time required by the particle to reach the centre of force from distance δ is δ^2/k.
12. A particle is at rest at a great height above the earth. It is then allowed to fall towards the earth. Neglecting air resistance show that it requires about $\frac{8}{11}$th of the total time of fall to traverse the first half of the distance.

13. Show that the gravitational potential due to a thin uniform circular disc at a point on its axis is

$$\Phi(z) = -2\pi G\sigma[\sqrt{z^2 + a^2} - |z|]$$

where σ is the mass per unit area of the disc and z is the distance of the point from the centre of the disc.

14. Calculate the potential due to a thin circular ring of very small radius a and mass M at a point in the plane of the ring but lying outside the ring ($r \gg a$).

15. Consider a body of any arbitrary shape and having very large mass and a spherical surface exterior to the body. Show that the average value of the potential due to the body taken over the spherical surface is equal to the value of the potential at the centre of the sphere.

16. The interaction energy between two dipole moments $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ with distance $\mathbf{r}$ is written as

$$V = \frac{3(\boldsymbol{\mu}_1 \cdot \mathbf{r})(\boldsymbol{\mu}_2 \cdot \mathbf{r})}{r^3} - \frac{\boldsymbol{\mu}_1 \cdot \boldsymbol{\mu}_2}{r^3}$$

Show that this is equivalent to

$$V = \boldsymbol{\mu}_1 \cdot \nabla\left(\boldsymbol{\mu}_2 \cdot \nabla \frac{1}{r}\right)$$

or $$V = \frac{\mu_1 \mu_2}{r^3}(2 \cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2 \sin \varphi)$$

where θ_1 is the angle between $\boldsymbol{\mu}_1$ and $\mathbf{r}$ and θ_2 and φ_2 are the polar and azimuthal angles of $\boldsymbol{\mu}_2$ with $\mathbf{r}$.

Use the result: $\cos \theta = \cos \theta_1 + \cos \theta_2 + \sin \theta_1 \sin \theta_2 \cos (\varphi_1 - \varphi_2)$.

17. In a non-rotating isolated mass, for example a star, the condition of equilibrium is

$$\nabla p + \rho \nabla \Phi = 0$$

where p is the total pressure, ρ the density, and Φ the gravitational potential. Show that at any given point the normals to the surfaces of constant density and constant gravitational potential are parallel.

18. An electric dipole of moment $\mathbf{d}$ placed at the origin produces an electric potential

$$\Phi(\mathbf{r}) = \frac{\mathbf{d} \cdot \mathbf{r}}{4\pi\epsilon_0 r^3}$$

Find the electric field $\mathbf{E}(\mathbf{r})$.

19. A particle of mass m moves under a central repulsive force mb/r^3 and is initially moving at distance a from the origin of the force with velocity V at right angles to a. Show that the equation of the path of the particle is

$$r \cos (p\theta) = a$$

where $p^2 = (b^2/a^2V^2) + 1$.

20. Two particles having masses m and M attract each other according to Newton's law of gravitation. Initially, they are at rest at an infinite distance apart. Show that their relative velocity of approach is

$$\sqrt{\frac{2G(M+m)}{a}}$$

when their separation is a.

21. Find the gravitational field and potential at any point z on the symmetry axis of a uniform solid hemisphere of radius a and mass M. The centre of the hemisphere is at $z = 0$.

22. Use Gauss' theorem to determine the gravitational field inside and outside a spherical shell of radius a, mass M and uniform density. Also calculate the resulting gravitational potential.

23. Find the gravitational field at distance x from an infinite plane sheet having surface density σ. Compare this result with the field just outside a spherical shell having the same surface density. What part of the total field is due to the immediately adjacent matter and what part due to more distant matter?

24. At what distance from the earth, on the line joining the earth to the sun, do the gravitational forces exerted on the mass by the earth and that by the sun become equal and opposite? Compare the result with the radius of the orbit of the moon around the earth. (Given: mass of the earth $= 5.96 \times 10^{24}$ kg, mass of the sun $= 1.97 \times 10^{30}$ kg, mean radius of the lunar orbit $= 3.84 \times 10^{8}$ m and mean distance between the sun and the earth $= 1.5 \times 10^{11}$ m).

5
Motion in a Central Force Field

A force is said to be a central force if it is always directed towards a fixed point. The forces exerted by two bodies on each other form the action and reaction pair and are equal and opposite according to Newton's third law of motion. We have seen that under mutual interaction, the two bodies move in such a way that their centre of mass remains fixed in space. The centre of mass is the centre of force in this case. The motion of a particle in a central force field is an important problem in physics, because this is the type of the motion performed by the planets around the sun, by satellites around the earth, by two charged particles with respect to each other and so on.

The motion of a particle in a central force field can be classified as: (i) *Bounded motion*: In this type of motion, the distance between two bodies never exceeds a finite limit. For example, the motion of planets around the sun, double star, etc. (ii) *Unbounded motion*: In this type of motion, the distance between the two bodies is infinite initially and finally. For example, scattering of alpha particles by nuclei of a gold foil as in the Rutherford experiment.

The motion of an electron around the nucleus, before the advent of quantum mechanics, was studied as a two-body problem in classical mechanics. The exact description of this motion, however, needs a quantum mechanical treatment.

It is always possible to reduce the motion of the two bodies to that of an equivalent single body in the central force field of mutual interaction. We shall discuss the *bounded central force field motion* in this chapter and the treatment of unbounded motion will be presented in Chapter 7.

The problem of two-body motion has an exact solution. The presence of the third body, however, complicates the situation and an exact

solution to the problem becomes an impossibility. Therefore one has to adopt the approximate methods to solve the three-body problems. We can always reduce many physical systems to a two-body problem either by neglecting the effects of the other bodies or by some other screening or averaging methods. For example, while studying the motion of a planet around the sun, the effect due to the presence of other planets is neglected. In this chapter, however, we shall restrict ourselves to two-body problems only.

5.1 EQUIVALENT ONE-BODY PROBLEM

Consider the motion of two particles of masses m_1 and m_2 separated by distance r. Let $\mathbf{F}^{ext}$ be the total external force acting on the system. Similarly, let $\mathbf{F}^{int}$ be the total internal force due to interaction between the two particles.

We can always write the total external force $\mathbf{F}^{ext}$ as a sum of external forces $\mathbf{F}_1^{ext}$ and $\mathbf{F}_2^{ext}$ acting separately on the two particles. Hence, we have

$$\mathbf{F}^{ext} = \mathbf{F}_1^{ext} + \mathbf{F}_2^{ext} \tag{5.1}$$

Further, by Newton's third law of motion

$$\mathbf{F}_{12}^{int} = -\mathbf{F}_{21}^{int} \tag{5.2}$$

these forces being the 'action and the reaction.'

The equations of motion of the two particles individually and as a system of two particles can be written as:

$$m_1\ddot{\mathbf{r}}_1 = \mathbf{F}_1^{ext} + \mathbf{F}_{12}^{int} \tag{5.3}$$

$$m_2\ddot{\mathbf{r}}_2 = \mathbf{F}_2^{ext} + \mathbf{F}_{21}^{int} \tag{5.4}$$

and

$$M\ddot{\mathbf{R}} = \mathbf{F}^{ext} \tag{5.5}$$

respectively (Fig. 5.1). Here $\mathbf{F}_{12}^{int}$ is the force exerted by particle of mass m_2 on the particle of mass m_1 and *vice versa*. In equation (5.5), we have put the total mass of the system

$$M = m_1 + m_2 \tag{5.6}$$

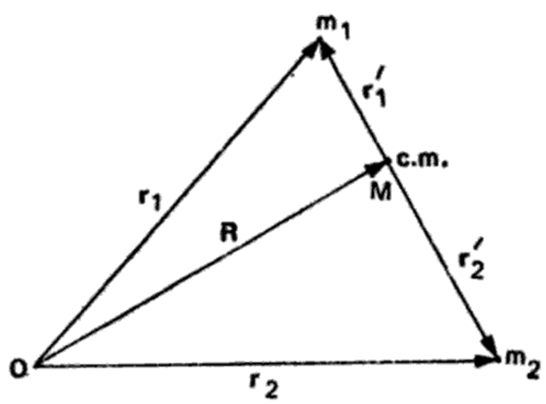


Fig. 5.1 Centre of mass of masses m_1 and m_2

and the position vector of the centre of mass of the system

$$\mathbf{R} = (m_1\mathbf{r}_1 + m_2\mathbf{r}_2)/(m_1 + m_2) \tag{5.7}$$

Let the position vector of particle 1 relative to particle 2 be

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \tag{5.8}$$

Solving equations (5.7) and (5.8) simultaneously for $\mathbf{r}_1$ and $\mathbf{r}_2$, we get

$$\mathbf{r}_1 = \mathbf{R} + \frac{m_2}{m_1 + m_2}\mathbf{r} \tag{5.9}$$

and

$$\mathbf{r}_2 = \mathbf{R} - \frac{m_1}{m_1 + m_2}\mathbf{r} \tag{5.10}$$

In order to write equations (5.3) and (5.4) in terms of $\mathbf{r}$, multiply equation (5.3) by m_2 and equation (5.4) by m_1 and subtract the latter from the former. Then, we get

$$m_1m_2(\ddot{\mathbf{r}}_1 - \ddot{\mathbf{r}}_2) = (m_2\mathbf{F}_{12}^{\text{int}} - m_1\mathbf{F}_{21}^{\text{int}}) + m_1m_2\left(\frac{\mathbf{F}_1^{\text{ext}}}{m_1} - \frac{\mathbf{F}_2^{\text{ext}}}{m_2}\right)$$

i.e.

$$m_1m_2\ddot{\mathbf{r}} = (m_1 + m_2)\mathbf{F}_{12}^{\text{int}} + m_1m_2\left(\frac{\mathbf{F}_1^{\text{ext}}}{m_1} - \frac{\mathbf{F}_2^{\text{ext}}}{m_2}\right) \tag{5.11}$$

Dividing equation (5.11) throughout by $m_1 + m_2$ and defining the reduced mass of the system by the formula

$$\mu = \frac{m_1m_2}{m_1 + m_2} \quad \text{or} \quad \frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2} \tag{5.12}$$

equation (5.11) becomes

$$\mu\ddot{\mathbf{r}} = \mathbf{F}_{12}^{\text{int}} + \mu\left(\frac{\mathbf{F}_1^{\text{ext}}}{m_1} - \frac{\mathbf{F}_2^{\text{ext}}}{m_2}\right) \tag{5.13}$$

(i) If no external force is acting

$$\mathbf{F}_1^{\text{ext}} = \mathbf{F}_2^{\text{ext}} = 0 \tag{5.14}$$

or (ii) if external forces $\mathbf{F}_1^{\text{ext}}$ and $\mathbf{F}_2^{\text{ext}}$ are proportional to the masses of the particles on which they act and produce equal accelerations in the two particles, i.e. if

$$\frac{\mathbf{F}_1^{\text{ext}}}{m_1} = \frac{\mathbf{F}_2^{\text{ext}}}{m_2} \tag{5.15}$$

equation (5.13) reduces to

$$\mu\ddot{\mathbf{r}} = \mathbf{F}_{12}^{\text{int}} \tag{5.16}$$

This is the equation of motion of a particle having mass equal to reduced mass μ and moving under the action of force $\mathbf{F}_{12}^{\text{int}}$. Such a reduction of a two-body problem to an equivalent one-body problem is very convenient and is very often used in classical as well as in quantum mechanics. *The reduction is equivalent to replacing the system by a mass equal to the reduced mass and considering the acceleration produced in it due to the interaction force.*

Equation (5.16) together with equation (5.5) represents the motion of a two-body system under the action of internal and external forces as long as the conditions mentioned in equations (5.14) or (5.15) are valid. The

condition of equation (5.15) is realised if the centre producing the external force is at a considerable distance from the system and the force due to it on any mass is proportional to that mass. The famous example of a force of this type is that of a gravitational force. In the discussion of motion of the moon around the earth, forces due to the sun, as an approximation, can be assumed to satisfy the condition of equation (5.15).

If the internal forces are of attraction and these are the only forces acting on the system, the two bodies move around the centre of mass which acts as the centre of force, i.e. a point towards which the forces are directed.

If the mass of one of the particles is extremely large as compared to that of the other, say $m_1 \gg m_2$, then the reduced mass is simply

$$\mu = \frac{m_1 m_2}{m_1 + m_2} = \frac{m_2}{1 + \dfrac{m_2}{m_1}}$$

or

$$\mu \simeq m_2 \text{ as } \frac{m_2}{m_1} \to 0$$

In this case, the centre of mass of the system coincides with the centre of mass of the heavier body of mass m_1. This approximation is equivalent to *neglecting the 'recoil' of* m_1. This is used in Bohr's theory of the hydrogen atom or in the case of motion of a satellite around the earth or that of the earth around the sun and so on. This fact can be easily understood if we apply Newton's second law of motion to these cases. Since mass m_2 is small, force $\mathbf{F}_{21}^{\text{int}}$ produces appreciable acceleration in m_2, whereas force $\mathbf{F}_{12}^{\text{int}}$ produces negligible acceleration in m_1. That is why *'an apple appears to fall towards the earth and not the earth towards the apple'*.

5.2 MOTION IN A CENTRAL FORCE FIELD

A force is said to be a central force if it acts along the position vector of a particle drawn from the centre of the force. Then, it can be expressed as

$$\mathbf{F}(\mathbf{r}) = \hat{\mathbf{e}}_r F(\mathbf{r}) \tag{5.17}$$

where $\hat{\mathbf{e}}_r$ is the unit vector along the direction of the position vector $\mathbf{r}$. If $\mathbf{F}(\mathbf{r})$ is positive, the force is repulsive and if $\mathbf{F}(\mathbf{r})$ is negative, the force is attractive. Examples of the attractive force are the gravitational forces between two mass points, the electrostatic force between two unlike charges, etc. An example of repulsive force is the electrostatic force between the like charges as in the case of scattering of alpha particles by nuclei, etc.

We now use the results of section 5.1 when only internal forces—either attractive or repulsive—are acting on the particles. Let the origin be the centre of force and the particle of reduced mass μ be at distance $\mathbf{r}$, the separation distance between the two particles. The motion of this hypothetical particle in the central force, which is the force of interaction

between the two particles, is equivalent to the two-body motion. Potential energy V is a function of the distance of this hypothetical particle from the centre of force only and does not depend upon the orientation. Hence the system has a *spherical symmetry.* As a result of this, the angular momentum of the particle is conserved. This is easily seen since the torque

$$\mathbf{N} = \frac{d\mathbf{L}}{dt} = \mathbf{r} \times \mathbf{F}(\mathbf{r}) = \mathbf{r} \times \hat{\mathbf{e}}_r F(\mathbf{r}) = 0$$

because $\mathbf{r} \times \hat{\mathbf{e}}_r = 0$. This gives

$$\mathbf{L} = \text{const}$$

or

$$\mathbf{r} \times \mathbf{p} = \text{const} \tag{5.18}$$

From equation (5.18), it is obvious that the plane containing $\mathbf{r}$ and $\mathbf{p}$ is always perpendicular to $\mathbf{L}$. Hence, the component of $\mathbf{L}$ along any axis through the centre of the force is constant. Further, since there is no component of $\mathbf{F}(\mathbf{r})$ perpendicular to the plane containing $\mathbf{r}$ and $\mathbf{p}$, the motion will always remain in the plane of $\mathbf{r}$ and $\mathbf{p}$ which is also the plane containing initial position and momentum.

We come across a special case when $\mathbf{L} = 0$. In this case, $\mathbf{r}$ is parallel to $\mathbf{p}$ and the motion is along a straight line.

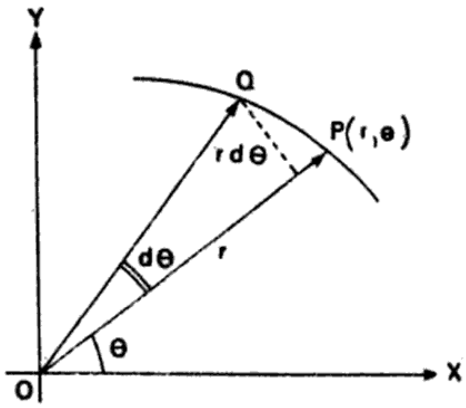


Fig. 5.2 Motion of centre of mass with respect to centre of force

Let us choose the x-, y-axes in the plane of motion with O as the origin. Let r and θ be the plane polar coordinates of a point where a particle of mass $\mu = \dfrac{m_1 m_2}{m_1 + m_2}$ is situated (Fig. 5.2). Now, the radial and the transverse accelerations are given by, (refer to equation 2.24)

$$a_r = \ddot{r} - r\dot{\theta}^2$$

and

$$a_\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta}$$

Hence, the equations of motion along r and θ directions are

$$\mu\ddot{r} - \mu r\dot{\theta}^2 = F(r) \tag{5.19}$$

and

$$\mu r\ddot{\theta} + 2\mu\dot{r}\dot{\theta} = 0 \tag{5.20}$$

respectively.

Equation (5.20) can be written as

$$\frac{d}{dt}(\mu r^2\dot{\theta}) = 0 \tag{5.21}$$

Hence $\qquad \mu r^2\dot{\theta} = \text{const}$

But $\qquad \mu r^2\dot{\theta} = L$, the angular momentum

Hence, we have

$$L = \mu r^2\dot{\theta} = \text{const} \tag{5.22}$$

Equation (5.22) can be interpreted in another way as follows: From Fig. 5.2, the area swept out by the radius in time dt is

$$dA = \tfrac{1}{2}r\cdot r\,d\theta$$

Hence, the rate at which this area is swept out or the areal velocity is given by

$$\dot{A} = \frac{dA}{dt} = \tfrac{1}{2}r^2\frac{d\theta}{dt} = \tfrac{1}{2}r^2\dot{\theta} = \frac{L}{2\mu} = \text{const} \tag{5.23}$$

by virtue of equation (5.22).

Or, we can write equation (5.20) in the form

$$\frac{d}{dt}(\tfrac{1}{2}r^2\dot{\theta}) = \frac{d}{dt}(\dot{A}) = 0, \text{ since } \mu \neq 0$$

This gives $\dot{A} = \text{const}$. Thus, the areal velocity of the particle in a central force field is constant. This is the proof of Kepler's second law of planetary motion which will be stated in article 5.6.

If the force is conservative, we get another integral of the motion, i.e. total energy E which is given by

$$\begin{aligned} E &= T + V \\ &= \tfrac{1}{2}\mu\dot{r}^2 + \tfrac{1}{2}\mu r^2\dot{\theta}^2 + V(r) \\ &= \tfrac{1}{2}\mu\dot{r}^2 + \frac{L^2}{2\mu r^2} + V(r) \end{aligned} \tag{5.24}$$

The fact that total energy E is a constant of motion can be proved by writing equation (5.19) in the form

$$\mu\ddot{r} = -\frac{d}{dr}\left(V + \frac{L^2}{2\mu r^2}\right) \tag{5.25}$$

where $F(r)$ is written as $-\dfrac{dV}{dr}$. Multiplying both sides of equation (5.25) by $\dot{r}$ and substituting $\dot{r}\dfrac{d}{dr} = \dfrac{dr}{dt}\dfrac{d}{dr} = \dfrac{d}{dt}$ on the right-hand side, we get

$$\mu\dot{r}\ddot{r} = -\frac{d}{dt}\left[V + \frac{L^2}{2\mu r^2}\right]$$

or

$$\frac{d}{dt}\left[\tfrac{1}{2}\mu\dot{r}^2 + \frac{L^2}{2\mu r^2} + V\right] = 0$$

Hence

$$\tfrac{1}{2}\mu\dot{r}^2 + \frac{L^2}{2\mu r^2} + V = \text{const}$$

Equation (5.24) can be solved for r as follows:
we have

$$\dot{r} = \frac{dr}{dt} = \left[\frac{2}{\mu}\left(E - V - \frac{L^2}{2\mu r^2}\right)\right]^{1/2} \tag{5.26}$$

or

$$\int_{r_0}^{r} \frac{dr}{\left[\frac{2}{\mu}\left(E - V - \frac{L^2}{2\mu r^2}\right)\right]^{1/2}} = \int_0^t dt = t \tag{5.27}$$

Equation (5.22) can be solved for θ. Thus

$$\dot{\theta} = \frac{d\theta}{dt} = \frac{L}{\mu r^2}$$

and on integration of

$$\int_{\theta_0}^{\theta} d\theta = \int_0^t \frac{L}{\mu r^2}\, dt$$

we get the solution

$$\theta = \theta_0 + \int_0^t \frac{L}{\mu r^2(t)}\, dt \tag{5.28}$$

In writing down the solutions in equations (5.27) and (5.28), we have used the conditions

at $t = 0$, $r = r_0$ and $t = t$, $r = r$

Similarly at $t = 0$, $\theta = \theta_0$ and $t = t$, $\theta = \theta$

Thus, the solutions of the equations of motion in principle may be obtained, i.e. the solutions of equations (5.19) and (5.20) in terms of four constants of integration, viz. L, E, θ_0 and r_0 which can be evaluated when the initial position and the velocity of the particle are known. These are not the only constants that can be used. We could have alternately chosen constants such as r_0, θ_0, $\dot{r}_0$ and $\dot{\theta}_0$. This set of constants is always expressible in terms of the set of constants we have chosen.

In the above method of obtaining solutions of the equation of motion we expressed both r and θ as functions of time. We are very often interested in the shape of the orbit, i.e. in the way in which r depends upon θ. To find this dependence, we eliminate time dependence by using equation (5.22) and write

$$d\theta = \frac{d\theta}{dt}\frac{dt}{dr}\, dr = \frac{\dot{\theta}}{\dot{r}}\, dr = \frac{L}{\mu r^2 \dot{r}}\, dr \tag{5.29}$$

This can be written as

$$d\theta = \frac{(L/r^2)\, dr}{\mu \dot{r}} \tag{5.30}$$

But, $\mu\dot{r} = \left[2\mu\left(E - V - \frac{L^2}{2\mu r^2}\right)\right]^{1/2}$, from equation (5.26). Substituting this in equation (5.30), we get

$$d\theta = \frac{(L/r^2)\, dr}{\left[2\mu\left(E - V - \frac{L^2}{2\mu r^2}\right)\right]^{1/2}}$$

Integrating this equation, we get

$$\theta(r) - \theta_0 = \int_{r_0}^{r} \frac{(L/r^2)\, dr}{\left[2\mu\left(E - V - \frac{L^2}{2\mu r^2}\right)\right]^{1/2}} \tag{5.31}$$

We have thus obtained the formal solution in the form of an integral which can be solved easily in some specific forms of the law of force. The most important law of force is the inverse square law and this case will be considered in article 5.4.

5.3 GENERAL FEATURES OF THE MOTION

Equivalent One-Dimensional Problem

If we express $\dot{\theta}$ in terms of the angular momentum, the equation of motion (5.19) in the radial direction can be written as

$$\mu \ddot{r} = F(r) + \frac{L^2}{\mu r^3} \tag{5.32}$$

The second term on the right-hand side of equation (5.32), viz. $L^2/\mu r^3$ is traditionally known as the centrifugal force. It is not strictly a 'force', since it does not arise out of interaction between the two particles. It arises due to accelerated motion of the particle, in the orbit and is a pseudo- or false-force. We shall discuss this type of force in details in Chapter 10. Because of its wide use we shall, however, use the term centrifugal force.

Equation (5.32) can be looked upon as a one-dimensional equation of motion in r with effective potential energy

$$V_e(r) = -\int\left[F(r) + \frac{L^2}{\mu r^3}\right] dr \tag{5.33}$$

or

$$V_e(r) = V(r) + \frac{L^2}{2\mu r^2} \tag{5.34}$$

Here we have used

$$F(r) = -\frac{dV(r)}{dr} \tag{5.35}$$

in obtaining equation (5.34). Further, the potential energy of the particle when situated at infinity, i.e. $V(\infty)$ is taken as zero. The term $L^2/2\mu r^2$ is the potential energy of the particle arising as a result of centrifugal force.

Now, the total energy of the particle in the central force field is the sum of its kinetic energy and the effective potential energy. Thus

$$E = \tfrac{1}{2}\mu \dot{r}^2 + V_e \tag{5.36}$$

Since, the system is conservative

$$E = T + V = \tfrac{1}{2}\mu \dot{r}^2 + V_e = \text{const} \tag{5.37}$$

Motion in Arbitrary Potential Field

We shall now discuss the motion of a particle in an arbitrary potential

field. For this, consider the one-dimensional motion in r with velocity, from equation (5.37)

$$\dot{r} = \sqrt{\frac{2}{\mu}(E - V_e)} \tag{5.38}$$

Let us suppose that the particle has total energy E as shown by the dotted line in Fig. 5.3. The arbitrary potential field is represented by the curve drawn in continuous line. It is seen from Fig. 5.3 that at points $r = r_1$ and r_2, straight line $E =$ constant intersects the potential energy curve.

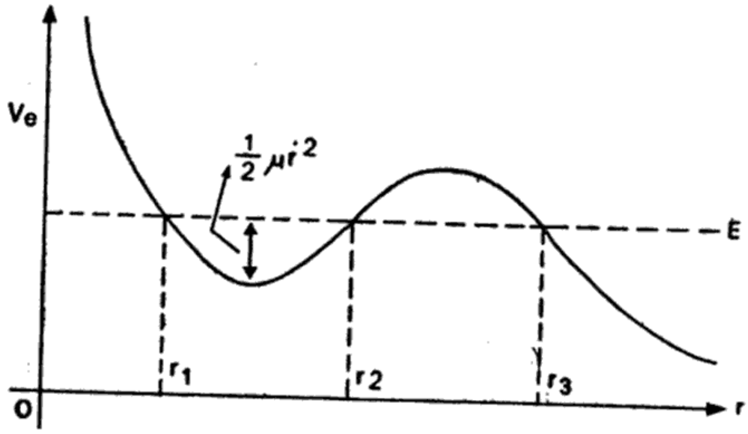


Fig. 5.3 Curve representing arbitrary potential field

These points correspond to $E = V_e$. Hence, at these points $\dot{r} = 0$ from equation (5.38). Such points, where the radial velocity is zero, are called the turning points. At all other points between r_1 and r_2 there exist certain differences between the values of E and V_e. This is represented by the ordinate between straight line $E =$ constant and the curve representing V This is obviously the kinetic energy of the particle at that point.

The whole range of value of r can now be divided into various regions as follows:

(i) ***Region for which*** $r < r_1$: In this region, the potential energy V_e is greater than total energy E. Hence, the kinetic energy will be negative and velocity $\dot{r}$ will be imaginary. Hence, this region is forbidden for the particle.

(ii) ***Region for which*** $r_1 \leqslant r \leqslant r_2$: In this region, the total energy E is greater than potential energy V_e. This region has $r = r_1$ and $r = r_2$ as the turning points. Distances r_1 and r_2 are called the *apsidal distances*. The motion of the particle is, therefore, oscillatory in the potential well and the particle will be confined to this region. The particle does not possess enough kinetic energy to cross over the potential barriers on either side. The orbit of the particle may not be

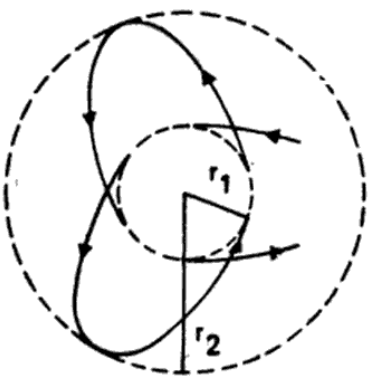


Fig. 5.4 Bounded moti· . bet-- ween circles of radii r_1 and r_2

closed and might be bounded in the region between two circles of radii r_1 and r_2 (Fig. 5.4).

(iii) *Region for which* $r_2 < r < r_3$: This region is also forbidden for the particle since total energy E is smaller than potential energy V_e.

(iv) *Region for which* $r \geqslant r_3$: This region has only one turning point, namely the one at $r = r_3$. A particle initially at an infinite distance from the centre of force is able to approach upto $r = r_3$ where it gets rebounded and proceeds unhindered to infinity again. Such a phenomenon is called *scattering* (Fig. 5.5).

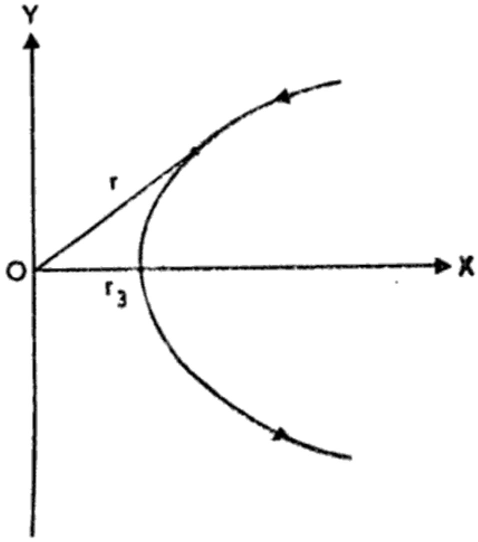


Fig. 5.5 The phenomenon of scattering

The motion is bounded corresponding to a situation as in region (ii) while it is unbounded corresponding to a situation as in region (iv).

If the total energy of a particle in a conservative force field is zero, i.e.

$$E = T + V_e = 0$$

or
$$T = -V_e \tag{5.39}$$

then the particle situated at any point will move to infinity. Thus, the motion would be unbounded for any value of r. The velocity of the particle as obtained from equation (5.39) is called the *escape velocity*.

The nature of the motion of the particle discussed earlier with the help of an arbitrary potential energy curve is extremely helpful in understanding the nature of the orbits. This above description of the motion is of the classical nature. The quantum mechanical nature of the motion is substantially different. In quantum mechanical description, firstly, the particle does not possess any arbitrary energy E in the potential well. The energy will have discrete values. Secondly, even if $E < V_e$, the particle always has some probability of penetrating the potential barrier. This is the so-called '*tunnelling effect*'. Thirdly, the particle inside the potential well will not have zero kinetic energy corresponding to the bottom position of the well but will have some finite minimum energy called '*zero point energy*'

5.4 MOTION IN AN INVERSE-SQUARE LAW FORCE FIELD

Let us now consider the case in which the force exerted by one particle on the other varies inversely as the square of the distance between them. Then

$$F(r) = \frac{k}{r^2}$$

where k is the constant of proportionality.

But, since $F(r) = -\frac{dV}{dr}$, i.e. $-\frac{dV}{dr} = \frac{k}{r^2}$, hence

$$V(r) = +\frac{k}{r} \tag{5.40}$$

Then the effective potential energy is given by

$$V_e = \frac{k}{r} + \frac{L^2}{2\mu r^2} \tag{5.41}$$

The value of constant k depends upon the nature of the physical problem. For example, if we are considering the gravitational force between the two spherical bodies having masses m_1 and m_2, then $k = -Gm_1m_2$, where G is the universal constant of gravitation and has a magnitude $G = 6{\cdot}67 \times 10^{-11}$ N-m^2/kg^2. The negative sign indicates that it is a force of attraction. In the case of electrostatic force between two charges q_1 and q_2, in free space, $k = \frac{q_1q_2}{4\pi\epsilon_0}$, where ϵ_0 is permittivity of free space. In this case the constant k is negative if the force is that of attraction and it is positive if the force is of repulsion. The nature of the orbit will depend upon the sign of k.

The graphs of effective potential energy V_e against the distance r are plotted for various values of k in Fig. 5.6. The dotted curves are for the cases for which the angular momentum is zero. It is clearly seen from Fig. 5.6 that for a repulsive force ($k \geqslant 0$), the potential energy curve does not have a minimum as in the case of a similar curve for a force of attraction ($k < 0$).

A particle having any arbitrary value of energy approaches the centre of force from infinity ($r = \infty$), reaches the closest distance of approach, turns around and again moves to infinity. The nature of orbits for such unbounded motion is sketched in Fig. 5.7 for positive and negative values of k. The distances of closest approach from centre O of the force are shown by the dotted lines in Fig. 5.7.

The orbit for $k = 0$ is just a straight line since no force is existing.

If the force is of attraction, the effective potential is

$$V_e = -\frac{|k|}{r} + \frac{L^2}{2\mu r^2} \tag{5.42}$$

In this case, the orbit of the particle is around the centre of force O. If the force is of repulsion, the particle follows an orbit along which it is

deflected away from the centre of the force and it moves towards infinity.

We now plot the graphs of two terms of the potential energy, viz. $-\frac{|k|}{r}$ and $\frac{L^2}{2\mu r^2}$ present on the right-hand side of equation (5.42) against distance r (dotted curves in Fig. 5.8). The variation of resultant or the

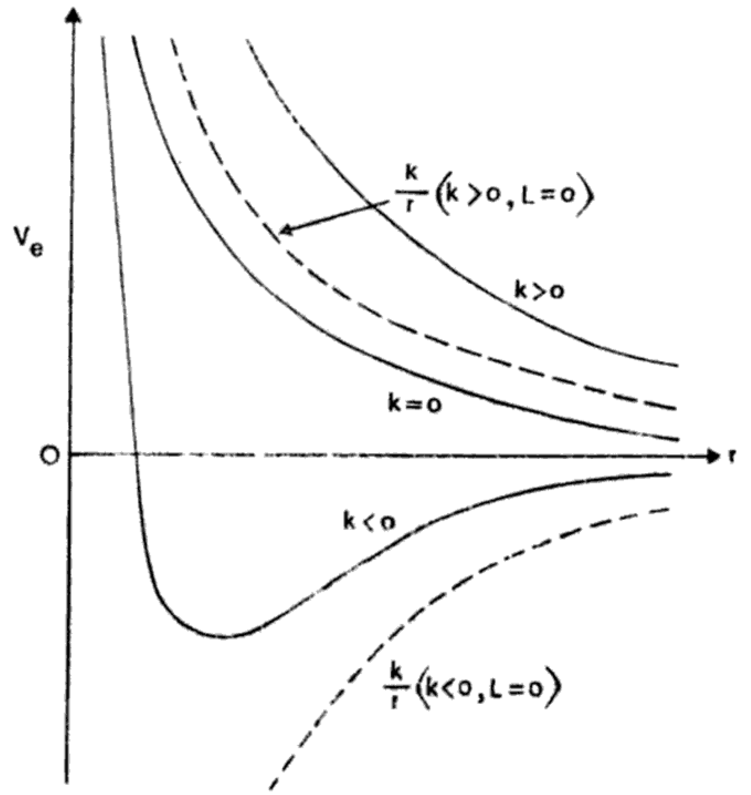


Fig. 5.6 Variation of effective potential V_e with distance r for potential k/r and positive, zero and negative values of k. For dotted curves $L = 0$ and line curves $L \neq 0$

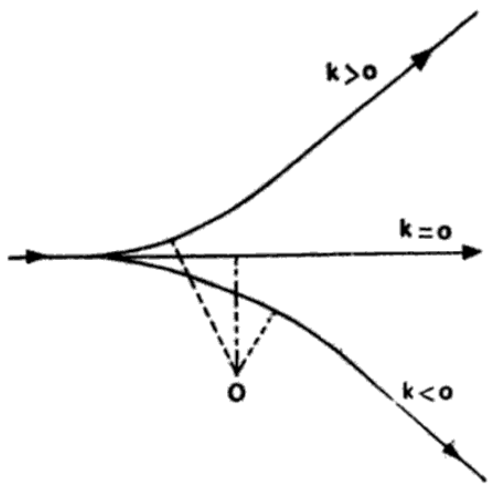


Fig. 5.7 Different types of orbits in the case of unbounded motion

effective potential V_e with respect to distance r is shown by a continuous curve. The range in which r can vary is $0 \leqslant r \leqslant \infty$.

Consider now a few special cases of total energies possessed by the particle (Fig. 5.8).

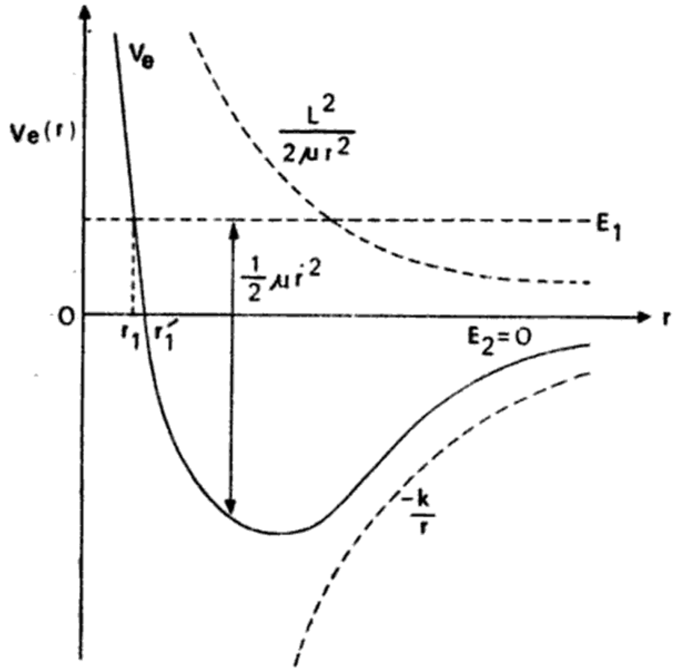


Fig. 5.8 Variation of effective potential V_e with radial distance r for inverse square law field. Dotted curves have $L = 0$

(i) For total energy E_1, we get the intersection of straight line $E_1 = \text{constant}$ with the potential energy curve at $r = r_1$. This value corresponds to the case when radial component of kinetic energy of the particle is zero. In this case, r_1 is a real root of the equation

$$E_1 = V_e = -\frac{|k|}{r} + \frac{L^2}{2\mu r^2} \tag{5.43}$$

The particle moves in such a way that it has only one turning point at $r = r_1$. Hence, this motion corresponds to scattering, i.e., an unbounded motion as mentioned above.

(ii) If the energy of the particle has value $E_2 = 0$, we again have a similar case. The two roots of equation (5.43) in this case are $r = r_1'$ and $r = \infty$. The particle goes to infinity but its radial velocity falls off continuously and becomes zero at infinity.

(iii) For energy $E_3 < 0$, the two roots r_2 and r_3 of equation (5.41) are real and distinct (Fig. 5.9). The motion of the particle is, therefore, bounded between these two values. The orbits are closed and elliptic. Distances r_2 and r_3 give the positions of the pericentre and the apocentre of the orbit respectively. These points are termed perihelion and aphelion, respectively, in the case of motion of a planet around the sun. The terms perigee and apogee are used to denote these points in the case

of motion of the earth.

(iv) For total energy E_4 shown in Fig. 5.9, the two roots of equation (5.42) coincide and the resulting orbit is circular. Since straight line $E_4 =$ constant is tangential to the potential energy curve, we must have

$$\frac{dV_e}{dr} = 0 \text{ at the point of contact}$$

i.e.

$$\frac{dV}{dr} - \frac{L^2}{\mu r^3} = 0$$

where V is the potential energy that satisfies the relation $F(r) = -\dfrac{dV}{dr}$

or

$$F(r) = -\frac{dV}{dr} = -\frac{L^2}{\mu r^3} = -\mu r\dot{\theta}^2 \tag{5.44}$$

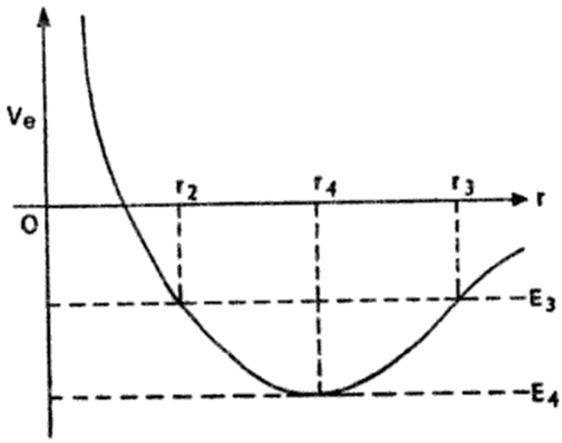


Fig. 5.9 Variation of V_e with r showing minimum (r_2) and maximum (r_3) radii of orbit for total energy E_3. For energy E_3, orbit is a circle of radius r_4

Thus, $F(r)$ is seen to be just equal to the centripetal force required to obtain circular motion of the particle around the centre of the force. Thus, $F(r)$ is the centripetal force that maintains the orbit.

5.5 EQUATION OF THE ORBIT

We know that the equation of motion of a particle subjected to a central force is

$$\mu\ddot{r} = F(r) + \frac{L^2}{\mu r^3}$$

We now put this equation in a form which is suitable to carry out the study of the nature of orbits. For this, we introduce a variable

$$u = \frac{1}{r}$$

Then, we can write

$$\frac{du}{d\theta} = -\frac{1}{r^2}\frac{dr}{d\theta}$$

$$= -\frac{1}{r^2}\frac{dr}{dt}\frac{dt}{d\theta}$$

But,

$$\frac{dr}{dt} = \dot{r} \text{ and } \frac{d\theta}{dt} = \dot{\theta}$$

Hence, we get

$$\frac{du}{d\theta} = \frac{-\dot{r}}{r^2\dot{\theta}}$$

$$= -\frac{\mu}{L}\dot{r}, \quad \text{by equation (5.22).}$$

Further

$$\frac{d^2u}{d\theta^2} = \frac{d}{d\theta}\left(\frac{du}{d\theta}\right)$$

$$= \frac{d}{d\theta}\left(-\frac{\mu}{L}\dot{r}\right)$$

$$= \frac{d}{dt}\left(-\frac{\mu}{L}\dot{r}\right)\frac{dt}{d\theta} = -\frac{\mu}{L}\frac{\ddot{r}}{\dot{\theta}}$$

$$= -\frac{\mu^2r^2}{L^2}\ddot{r}, \quad \text{by equation (5.22).}$$

Substituting these values in equation (5.32), we get

$$\frac{d^2u}{d\theta^2} + u = \frac{-\mu}{L^2u^2}F\left(\frac{1}{u}\right) \tag{5.45}$$

If the force-field obeys the inverse square law, then

$$F(r) = \frac{|k|}{r^2}$$

or

$$F\left(\frac{1}{u}\right) = |k|u^2$$

With this, equation (5.45) reduces to

$$\frac{d^2u}{d\theta^2} + u = \frac{|k|\mu}{L^2} \tag{5.46}$$

We now introduce variable

$$y = u - \frac{|k|\mu}{L^2}$$

Then, equation (5.46) assumes the form

$$\frac{d^2y}{d\theta^2} + y = 0 \tag{5.47}$$

This equation is a second-order differential equation in y as a function of θ and has a solution

$$y = A\cos(\theta - \theta_0) \tag{5.48}$$

where A and θ_0 are the constants of integration. Constant θ_0 is the value of variable θ when y assumes the maximum value.

Substituting value of y back in terms of r we get

$$\frac{1}{r} = \frac{|k|\mu}{L^2} + A\cos(\theta - \theta_0)$$

or
$$\frac{1}{r} = \frac{|k|\mu}{L^2}\left[1 + \frac{L^2 A}{|k|\mu}\cos(\theta - \theta_0)\right] \tag{5.49}$$

or
$$\frac{L^2/|k|\mu}{r} = 1 + \frac{L^2 A}{|k|\mu}\cos(\theta - \theta_0) \tag{5.50}$$

This last equation is similar to the equation

$$\frac{l}{r} = 1 + \epsilon \cos\theta$$

which is the equation of a conic section with the origin at the focus. Comparing equation (5.50) with this standard equation, we have

$$\text{semi-latus rectum, } l = \frac{L^2}{|k|\mu}$$

and
$$\text{eccentricity, } \epsilon = \frac{L^2 A}{|k|\mu}$$

The same result can be obtained by integrating equation (5.31), viz.

$$\theta(r) - \theta_0 = \int \frac{(L/r^2)\, dr}{\sqrt{2\mu\left(E - V - \dfrac{L^2}{2\mu r^2}\right)}}$$

with the substitution $V = -\dfrac{|k|}{r}$. The equation then assumes the form

$$\theta - \theta_0 = -\int \frac{du}{\sqrt{\dfrac{2\mu E}{L^2} + \dfrac{2\mu k}{L^2}u - u^2}} \tag{5.51}$$

when expressed in terms of variable $u = \dfrac{1}{r}$.

To integrate the right-hand side of equation (5.51), we can use the standard result

$$\int \frac{dx}{\sqrt{a + bx + cx^2}} = \frac{1}{\sqrt{-c}}\cos^{-1}\left(-\frac{b + 2cx}{\sqrt{b^2 - 4ac}}\right) \tag{5.52}$$

and get the same solution as in equation (5.50).

Since the cosine of an angle varies between $+1$ and -1, the turning points are given by

$$\left.\begin{aligned} \frac{1}{r_1} &= \frac{\mu|k|}{L^2} + A \\ \text{and} \qquad \frac{1}{r_2} &= \frac{\mu|k|}{L^2} - A \end{aligned}\right\} \tag{5.53}$$

Thus the value of constant A cannot exceed $\dfrac{\mu|k|}{L^2}$, otherwise it would make r_2 negative and hence meaningless.

The turning points are also the roots of the equation

$$E - V_e(r) = E + \frac{|k|}{r} - \frac{L^2}{2\mu r^2} = 0$$

This equation is a quadratic in $\frac{1}{r}\cdot$ Hence, its roots can be written as

$$\frac{1}{r_{1,2}}=\frac{\mu|k|}{L^2}\pm\frac{\mu|k|}{L^2}\sqrt{1+\frac{2EL^2}{\mu k^2}} \tag{5.54}$$

Comparing equations (5.53) and (5.54), we get

$$A=\frac{\mu|k|}{L^2}\sqrt{1+\frac{2EL^2}{\mu k^2}} \tag{5.55}$$

The eccentricity is then given by

$$\epsilon=\frac{L^2A}{\mu|k|}=\frac{L^2}{\mu|k|}\frac{\mu|k|}{L^2}\sqrt{1+\frac{2EL^2}{\mu k^2}}=\sqrt{1+\frac{2EL^2}{\mu k^2}} \tag{5.56}$$

(a) Nature of the Orbits

The nature of the orbit is entirely determined by the value of eccentricity ϵ of the orbit

$$\epsilon=\sqrt{1+\frac{2EL^2}{\mu k^2}}$$

The value of eccentricity ϵ depends upon total energy E. Following cases are possible depending upon the value of E.

Value of energy	Value of eccentricity	Nature of the orbit
$E>0$	$\epsilon>1$	hyperbola
$E=0$	$\epsilon=1$	parabola
$V_{e_{min}}<E<0$	$0<\epsilon<1$	ellipse
$E=V_{e_{min}}$	$\epsilon=0$	circle

For simplicity, we can always set $\theta_0=0$ and write equation (5.50) as

$$\frac{1}{r}=c(1+\epsilon\cos\theta) \tag{5.57}$$

where $$c=\frac{\mu|k|}{L^2}$$

(b) Elliptic Orbits

The ellipse is a curve traced out by a particle moving in such a way that the sum of its distances from two fixed foci O and O' is always constant.

Thus, $$OP+O'P=\text{const.}$$

or $$r+r'=2a \tag{5.58}$$

where a is the semi-major axis of the ellipse (Fig. 5.10).

The distance between the turning points is given by

$$r_1+r_2=2a \tag{5.59}$$

From equation (5.57), we can write

$$r_1=\frac{1}{c(1+\epsilon)}$$

and $$r_2=\frac{1}{c(1-\epsilon)}$$

Thus $$r_1+r_2=\frac{2}{c(1-\epsilon^2)} \tag{5.60}$$

Comparing equations (5.59) and (5.60), we get

$$a = \frac{1}{c(1-\epsilon^2)} \quad \text{or} \quad c = \frac{1}{a(1-\epsilon^2)} \tag{5.61}$$

Substituting this value of c in equation (5.57), we get

$$r = \frac{a(1-\epsilon^2)}{1+\epsilon\cos\theta} \tag{5.62}$$

This is the polar equation of the orbit.

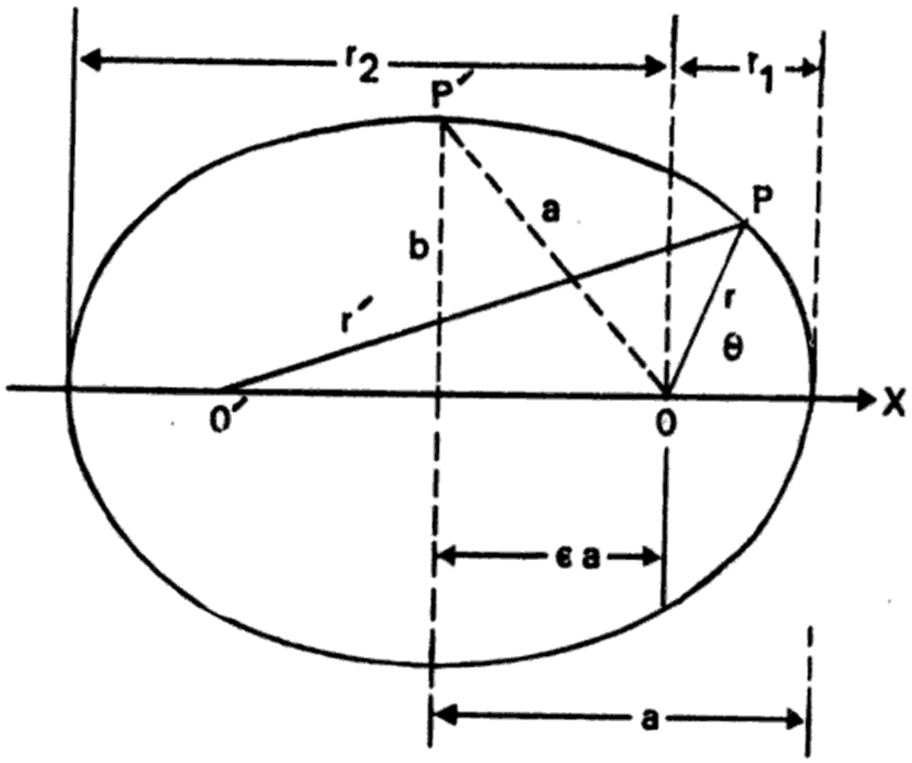


Fig. 5.10 Ellipse

From Fig. 5.10, the distance between the two foci is

$$OO' = r_2 - r_1 = \frac{2\epsilon}{c(1-\epsilon^2)} = 2a\epsilon \tag{5.63}$$

If the particle is situated at P', such that $OP' = O'P'$, its distance from O or O' must be equal to a since $OP' + O'P' = 2a$. Then, from Fig. 5.10, the semi-minor axis b is given by

$$b^2 = a^2 - a^2\epsilon^2$$
$$= a^2(1-\epsilon^2)$$

or

$$b = a\sqrt{(1-\epsilon^2)} \tag{5.64}$$

The ellipse will be converted into a circle when $\epsilon = 0$. But

$$\epsilon = \sqrt{1 + \frac{2EL^2}{\mu k^2}}$$

Hence, when $\epsilon = 0$, energy E of the particle is given by

$$E = \frac{\mu k^2}{2L^2} \tag{5.65}$$

The radius of the resulting circle is given by

$$a = \frac{1}{c} = \frac{L^2}{\mu\,|k|} = \frac{|k|}{2E} \tag{5.66}$$

(c) Hyperbolic Orbits

The hyperbola is a curve traced by a particle moving in such a way that the difference between its distances from the two foci O and O' is constant. Thus

$$\begin{aligned} O'P - OP &= \text{const} \\ \text{or} \qquad r' - r &= \pm 2a \end{aligned} \tag{5.67}$$

where $2a$ is the distance between the two vertices of the hyperbola (Fig. 5.11). The positive and negative signs on the right-hand side of equation (5.67) are used to describe the right or the left branch of the hyperbola respectively.

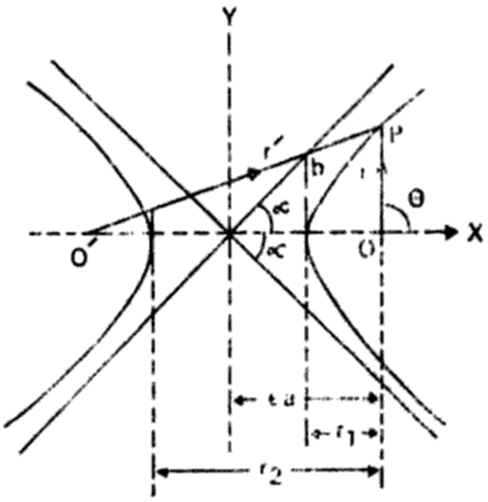


Fig. 5.11 Hyperbola

For a hyperbolic orbit

$$E > 0 \quad \text{and} \quad \epsilon > 1$$

Moreover, distance OO' between the two foci is given by

$OO' = 2a\epsilon$, as in the case of elliptic orbit.

The polar equation of this orbit is

$$r = \frac{a(\epsilon^2 - 1)}{\pm 1 + \epsilon \cos\theta} \tag{5.68}$$

In equation (5.68), the positive sign refers to the branch on the right and the negative sign to that on the left side. The asymptotes to the hyperbola make angle α with the x-axis.

This is the value of variable θ when r tends to infinity.

Hence,

$$\cos\alpha = \pm\frac{1}{\epsilon} \tag{5.69}$$

(d) Parabolic Orbits

A parabola is a curve traced by a particle in such a way that its distance from a fixed line (called a directrix) is equal to its distance from a fixed focus.

From Fig. 5.12, we can easily verify that

$$r = \frac{a}{1 - \cos\theta} \tag{5.70}$$

This is the polar equation of the parabola.

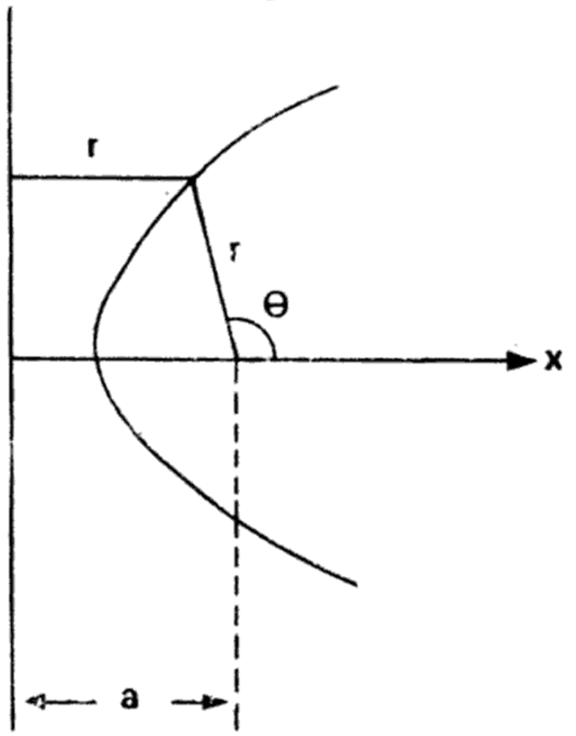


Fig. 5.12 Parabola

5.6 KEPLER'S LAWS OF PLANETARY MOTION

Kepler had announced the three laws of planetary motion based on the observations before Newton's laws of motion were formulated. Kepler's laws of planetary motion are:

(i) *The planet moves around the sun in an elliptic orbit with the sun at one of the foci of the orbit.*

(ii) *The areas swept out by the radius vector drawn from the centre of the sun to the centre of the planet in equal intervals of time are equal.* In other words, *the areal velocity is a constant.*

(iii) *The square of the period of revolution of the planet is directly proportional to the cube of the semi-major axis of the orbit.*

We have already pointed out that under the action of a force of attraction that obeys an inverse square law, the motion of the particle is a bounded motion in elliptic orbit. The force in the case of planetary motion is that exerted by the sun on the planet. The conservation of angular momentum in the central force-field motion leads us to Kepler's second law as discussed in article 5.1.

We now give the proof of Kepler's third law of motion. We know that the period of revolution is given by

$$\tau = \frac{\text{area of the elliptic orbit}}{\text{areal velocity}}$$

$$= \frac{\pi ab}{L/2\mu} = \frac{2\mu\pi a^2}{L}\sqrt{1-\epsilon^2} \tag{5.71}$$

But, the semi-major axis is given by

$$a = \frac{L^2}{\mu k}\frac{1}{1-\epsilon^2} = \left|\frac{k}{2E}\right| \tag{5.72}$$

Squaring equation (5.71) and substituting for a from equation (5.72), we get

$$\tau^2 = 4\pi^2\left|\frac{\mu}{k}\right| a^3 \tag{5.73}$$

This proves Kepler's third law.

If mass m of the planet is very small compared to mass M of the sun, we have

$$\mu = m$$

and $$k = GmM$$

Hence, we get

$$\tau^2 = \frac{4\pi^2}{GM}a^3 \tag{5.74}$$

It is clear from equation (5.74) that τ^2/a^3 is a constant for all the planets. We can also use equation (5.74) to find the mass of the sun.

QUESTIONS

1. How does a two-body problem reduce to a one-body problem? Compare the corresponding factors such as mass, distance and centre of mass in the two cases.
2. Explain what Kepler's second law implies about the force acting on a planet.
3. Discuss the developments of Kepler's law of planetary motion.
4. Explain what Kepler's first law implies when it is coupled with the second law.

5. What information about the force law comes from Kepler's third law?
6. Why is it difficult to determine the gravitational constant G accurately?
7. If the force $\mathbf{F}$ acting on the particle is inversely proportional to the cube of its distance from the centre of the force, what is the form of its trajectory?
8. What kind of problems can be solved using Gauss' law?
9. If the gravitational force acts on all bodies in proportion to their masses, why doesn't a heavy body fall faster than a light body?
10. How does the weight of a body vary en route from earth to moon? Would its mass vary?
11. Would we have more sugar to the kilogramme at pole or equator?
12. The earth is an oblate spheroid because of the flattening effect of the earth's rotation. Is a degree of latitude larger or smaller near either pole than near the equator? Why?
13. Is the central force whose magnitude is a function of the distance from the force centre a conservative force?
14. Explain that the orbital angular momentum is a constant of the motion governed by a central force whatever might be the law of force.
15. Explain the terms: escape velocity and tunnelling effect.
16. The nature of the orbit is determined by the value of its eccentricity $\epsilon = \sqrt{1 + \frac{2EL^2}{\mu k^2}}$. Discuss the various special cases depending upon the value of E and hence of ϵ.

PROBLEMS

1. Find the period of revolution of Mars given that the major axis of Mars is 1.5237 times that of the earth.
2. How far will a body, having acceleration g m/s^2, have to fall from rest to reach the same speed as that of a body falling freely from infinite distance to the surface of the earth in the earth's gravitational field?
3. Find the acceleration of a body falling freely through a tunnel imagined to be drilled through the centre of the earth, assuming that the density of earth is constant. Also find how distance r from the centre of the earth changes.
4. In problem 3, what would be the changes if the tunnel does not pass

through the centre of the earth?

5. While time t is negative, the path of a particle is given by the equations

$$r = -at \quad \text{and} \quad \varphi = -\frac{b}{t}$$

in the spherical coordinates where a and b are constants. What is the force law?

6. Suppose that a particle is attracted towards the z-axis by a force directly proportional to the square of its distance from the x-y plane and inversely proportional to its distance from the z-axis. Assuming that potential $V(r)$ exists for the particle, find the potential. Also find the additional perpendicular force that exists.

7. The periodic time of Venus is 224.7 days and that of earth 365.26 days. Find the ratio of the major axis of orbits of Venus and the earth.

8. A particle moves along circle

$$r = A \cos \varphi$$

Find the force law which the particle experiences.

9. A particle of mass m performs a uniform circular motion along the circumference of a circle of radius a. Express, in terms of its orbital speed, the minimum speed it must have or must be given in order to escape to infinity from a point on this circle.

10. A particle moves along an ellipse under the action of a central force. (a) If the centre of the ellipse is the force centre, show that $\mathbf{F} = -k\mathbf{r}$. (b) If the force centre is at one of the foci, show that $\mathbf{F} = -k\mathbf{r}/r^3$.

11. A central attractive force varies as r^m. The velocity of a particle in a circular orbit of radius r is twice the escape velocity from the same radius. Find m.

12. Equations of the orbits of a particle under the action of central forces are given by

    (i) $r = a(1 + \cos \theta)$

    (ii) $r = ae^{b\theta}$

    (iii) $\dfrac{1}{r} = a \cosh b(\theta - \theta_0)$

    Find the corresponding forces.

13. A double star is formed of two components, each having a mass equal to mass of the sun. The distance between them is the same as that between the earth and the sun. What is its orbital period?

14. A particle is moving in an inverse square field. Find $\mathbf{P}$ such that

$\mathbf{R} = \mathbf{L} \times \mathbf{P} - \dfrac{\mathbf{r}}{r}$ is a constant. What is the significance of $\mathbf{R}$ for trajectories of different types ?

15. A particle moves in a circular orbit under the action of an attractive force directed towards a fixed point on the circle. Show that the magnitude of the force varies inversely as the fifth power of the distance.

16. A planet moving round the sun is suddenly stopped. Find the time taken by the planet to fall into the sun. Express the time in terms of the period of revolution of the planet round the sun.

17. A comet travels in a parabolic orbit in the same plane as that of the earth's orbit. The distance of the point of closest approach of the comet from the sun is $\frac{1}{3}R$, where R is the radius of the earth's orbit around the sun, which is assumed to be circular. Show that the comet remains within the orbit of the earth for 74.5 days.

18. Kepler's laws of planetary motion follow from Newton's laws of motion and the law of gravitation. Historically, Newton deduced the law of gravitation from Kepler's laws. Starting from Kepler's laws, obtain the law of gravitation.

19. A particle moves along the parabola $y^2 = 4a^2 - 4ax$ where a is constant. The speed v of the particle is constant. Find components of its velocity and acceleration in rectangular and polar coordinates. Show that the equation of the parabola in polar coordinates is

$$r \cos^2 \frac{\theta}{2} = a$$

20. Discuss the types of motion that can occur under the action of a central force

$$f(r) = -\frac{k}{r^2} + \frac{k'}{r^3}$$

where $k > 0$ and k' may be greater or less than zero.

21. The Yukawa potential is given by

$$V(r) = \frac{ke^{-ar}}{r}, \; k < 0$$

Find the force and compare it with an inverse square law of force. Also discuss the types of motion of a particle of mass m under the action of such a force.

22. A satellite moves around the earth in an orbit which passes across the poles. The time at which it crosses each parallel of latitude is measured so that function $\theta(t)$ is known. Explain how you will find the perigee, the semi-major axis and the eccentricity of its orbit in terms of $\theta(t)$ and the value of g at the surface of the earth. Assume the earth to be a sphere of uniform density.

23. A particle moves in a circular orbit under the action of force $f(r) = -k/r^2$. If k is suddenly reduced to half its original value, show that the particle now moves along a parabola.

24. Show that the product of maximum and minimum linear speeds of a particle moving in an elliptic orbit is $(2\pi a/\tau)^2$, where τ is the period and a is the semi-major axis.

25. Consider the family of orbits in a central potential for which the total energy is constant. Show that, if a stable circular orbit exists, the angular momentum associated with this orbit is larger than that for any other orbit of the family.

6
Oscillations

We come across a large number of phenomena in nature in which a periodic motion is taking place. A periodic motion is one which repeats itself after a definite interval of time. For example, human pulse or heart beats, the pendulum clock, the occurrence of day and night, the motion of a mass attached to a spiral spring, atoms in a molecule or in a solid lattice and so on. It is observed that the displacement of a particle in periodic motion can be expressed in terms of sines and cosines of angles dependent on time.

If the particle performing a periodic motion moves to and fro over the same path, its motion is said to be oscillatory and the particle is called *oscillator*. If the maximum displacement of the oscillating particle from its equilibrium position, i.e. the amplitude of oscillation remains the same, the motion is called undamped motion. In practice, we observe that the amplitude of oscillation of a simple pendulum or that of a vibrating spiral spring goes on decreasing gradually, and ultimately the oscillation dies down. Such a motion is called a damped motion. This decrease in amplitude is due to some kind of resistive force offered by the medium in which oscillations are taking place. The energy of the oscillator is spent in overcoming this resistive force and goes on decreasing. The damping effect of the resistive force can be annuled by supplying energy to the oscillator to compensate for the energy dissipated by the resistive forces. The main spring of the watch or the clock supplies the necessary energy and the motion of the balance wheel or the pendulum is as if it were undamped.

Many times, the periodic motion of a system is maintained by a periodic driving force. Oscillations of such a system are called forced oscillations. For example, oscillation of air column in a resonator when a tuning fork is held at its mouth. The amplitude or the maximum displacement of the system, i.e., oscillator depends upon the frequency of the

driving force. There exists some frequency of the driving force at which the amplitude of oscillations becomes very large. When the two frequencies match, the driving system and the driven system are said to be in resonance. The phenomenon of resonance plays an important role in mechanical vibrations, in musical instruments, in alternating current circuits, etc.

In yet another case, one oscillating system may be set in oscillation by coupling it with another oscillating system. The oscillations of the system are then called the coupled oscillations

Because of the importance of oscillations in practically all branches of physics and engineering, we shall consider this phenomenon in detail.

In any periodic motion, the time required by the moving system to complete one round trip and occupy the same state of motion (i.e. phase) is called the period (τ) of the motion. Each round trip is called one oscillation or a cycle. The number of oscillations per unit time (i.e. per second) is called the frequency (ν) of the oscillation. It is obvious that

$$\nu = \frac{1}{\tau} \quad \text{or} \quad \tau = \frac{1}{\nu} \tag{6.1}$$

The S.I. unit of frequency is cycles per second (cps) or **hertz (Hz).**

If oscillations of a particle are damped, the particle will come to a stop at a point known as equilibrium position. In this position, the particle does not experience any force or the external force vanishes. The force acting on the particle is usually directed towards this position which is taken as the reference position for the measurement of displacement of the particle. Thus, displacement of the oscillating particle at any instant is the displacement of the oscillating particle from the equilibrium position at that instant.

Displacement of the oscillator measured from the equilibrium position changes periodically (both in magnitude and direction). Hence, its velocity and acceleration also change periodically (both in magnitude and direction). Consequently, by Newton's second law of motion, force must also be changing in magnitude and direction. This force, as we shall presently see in some examples, is usually directed towards the equilibrium position of the particle and hence it is often called the restoring force. This force is a function of the displacement in the case of oscillatory motion. Likewise the potential is also a function of the position of oscillator.

6.1 SIMPLE HARMONIC OSCILLATOR

We have mentioned above that for any periodic motion, there exists a position of the particle in which the net force acting on the particle is zero. This is the equilibrium position. Let $x = x_0$ denote this position. We now consider the one-dimensional periodic motion of the particle about this position. If we assume that the displacement is small, the potential energy function can be approximately taken to be equal to the first term

(that does not vanish) in its Taylor series expansion about the equilibrium position, viz. $x = x_0$.

The Taylor expansion of a function $U(x)$ is given by

$$U(x) = \sum_{n=0}^{\infty} \frac{1}{n!} U^{(n)}(x_0)[x - x_0]^n \tag{6.2}$$

where
$$U^n(x_0) = \left.\frac{d^n f(x)}{dx^n}\right|_{x=x_0}$$

Using this equation and remembering that $\frac{dU}{dx} = 0$ corresponding to minimum at $x = x_0$, we can write the potential energy as

$$U(x) = U(x_0) + \frac{1}{2!} U^{(2)}(x_0)(x - x_0)^2 + \ldots \tag{6.3}$$

But, $U(x_0)$ is the potential energy in the equilibrium position and can be chosen, although arbitrarily, to be equal to zero.

Thus
$$U(x_0) = 0$$

Then
$$U(x) = \tfrac{1}{2}U^{(2)}(x_0)(x - x_0)^2$$

or
$$U(x) = \tfrac{1}{2}k(x - x_0)^2 \tag{6.4}$$

where $k = U^{(2)}(x_0)$.

Since, for stable equilibrium, the potential energy in that position must be minimum, $U^{(2)}(x_0)$ must be positive, and hence k is also positive.

We further simplify the problem by shifting the origin of the co-ordinate system to the equilibrium point, viz. $x_0 = 0$.

Hence, from equation (6.4), we get

$$U(x) = \tfrac{1}{2}kx^2 \tag{6.5}$$

While obtaining this result we have neglected x^3 and the higher order terms in equation (6.3). This amounts to approximating the potential energy curve with the parabola given by equation (6.5).

With this approximation, the force acting on the particle at any point x is given by

$$F = -\frac{dU}{dx} = -kx \tag{6.6}$$

But, by Newton's second law of motion, the equation of motion of the particle is given by

$$m\ddot{x} = -kx \tag{6.7}$$

Constant k which is numerically equal to $\frac{m\ddot{x}}{x}$ is the force acting on the particle per unit displacement. This is called the force constant of the periodic motion of the particle. Equation (6.7) has been obtained after several simplifications of the general equation (6.3). The motion of the particle represented by equation (6.7) is, therefore, referred to as a simple harmonic motion. If higher order terms in equation (6.3) are used, the

force will no longer be proportional to displacement x and higher order terms will appear in the equation of motion (6.7). The motion of the particle will no longer be harmonic.

To solve equation (6.7), we rewrite it as

$$\ddot{x} = -\omega_0^2 x \tag{6.8}$$

where $\omega_0^2 = \dfrac{k}{m}$. This has to be integrated twice to obtain a solution for x. For the first integration the left-hand side of equation (6.8) can be expressed as

$$\ddot{x} = \frac{d}{dt}(\dot{x}) = \frac{d}{dx}(\dot{x})\frac{dx}{dt} = \frac{d\dot{x}}{dx}\dot{x}$$

Substituting this in equation (6.8), we get

$$\dot{x}\frac{d\dot{x}}{dx} = -\omega_0^2 x \tag{6.9}$$

Integrating equation (6.9) with respect to x, we get

$$\frac{\dot{x}^2}{2} = -\frac{\omega_0^2 x^2}{2} + c \tag{6.10}$$

where c is a constant of integration, and can be evaluated by using boundary conditions. Thus, when x is maximum, i.e., when $x = A$, velocity $\dot{x} = 0$. This initial condition in equation (6.10) gives

$$0 = -\frac{\omega_0^2 A^2}{2} + c$$

or

$$c = \frac{\omega_0^2 A^2}{2} \tag{6.11}$$

where A is amplitude of the oscillations.

Substituting this value of c in equation (6.10) and simplifying it we get

$$\dot{x}^2 = \omega_0^2(A^2 - x^2)$$

Or the velocity of the particle in terms of its displacement is

$$\dot{x} = \pm\,\omega_0\sqrt{A^2 - x^2} \tag{6.12}$$

In order to obtain displacement, we write equation (6.12) in the form

$$\frac{dx}{\pm\sqrt{A^2 - x^2}} = \omega_0\,dt$$

Integration can be easily performed by substituting $x = A\sin\theta$ and getting

$$d\theta = \omega_0\,dt \tag{6.13}$$

With the conditions, when

$$t = t_0,\ \theta = \theta_0,\ \text{where } \theta_0 = \sin^{-1}\frac{x_0}{A}$$

and

$$t = t,\ \ \theta = \theta,\ \ \text{where } \theta = \sin^{-1}\frac{x}{A}$$

integration of equation (6.13) yields

$$\theta - \theta_0 = \omega_0(t - t_0)$$

i.e.

$$\sin^{-1}\frac{x}{A} - \sin^{-1}\frac{x_0}{A} = \omega_0(t - t_0)$$

or

$$x = A\sin[\omega_0(t - t_0) + \theta_0] \tag{6.14a}$$

where $\theta_0 = \sin^{-1}\frac{x_0}{A}$ is known as the initial phase.

Equation (6.14a) gives the displacement of the particle performing simple harmonic motion. It reveals that the motion of the particle is a sinusoidal oscillation and has period

$$\tau_0 = \frac{2\pi}{\omega_0} = 2\pi\sqrt{\frac{m}{k}} \tag{6.15}$$

and frequency

$$\nu_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \tag{6.16}$$

This is called the natural frequency of the simple harmonic motion.

Equation (6.14a) can be simplified further by taking initial conditions as follows: at $t = t_0 = 0$, $x = x_0 = 0$. Thus, we start measuring the time when the particle is in the equilibrium position.

Then, $\theta_0 = \sin^{-1}\frac{x_0}{A} = 0$. Equations (6.14a) then reduce to

$$x = A\sin\omega_0 t \tag{6.14b}$$

We could also take the initial condition as at $t = t_0 = 0$, $x = x_0 = A$. Thus, we start measuring the time when the particle has the greatest displacement (= amplitude). Then

$$\theta_0 = \sin^{-1}\frac{x_0}{A} = \sin^{-1}\frac{A}{A} = \sin^{-1}1 = \frac{\pi}{2}$$

and

$$x = A\sin\left(\omega_0 t + \frac{\pi}{2}\right)$$

or

$$x = A\cos\omega_0 t \tag{6.14c}$$

The simple harmonic motion discussed in this article is undamped since the only force acting on the particle is the restoring force and no other resistive force has been taken into consideration. The equations become somewhat different when a resistive force or a damping force is taken into consideration. This is considered in article (6.2).

(a) Energy of a Simple Harmonic Oscillator

The kinetic energy of an oscillator at any instant is given by $\frac{1}{2}m\dot{x}^2$, where $\dot{x}$ is its velocity at that instant.

The kinetic energy, by using equation (6.14), is given by

$$T = \tfrac{1}{2}m\dot{x}^2 = \tfrac{1}{2}mA^2\omega_0^2\cos^2[\omega_0(t - t_0) + \theta_0]$$

From this, we see that the maximum kinetic energy of the particle is

$$T_{max} = \tfrac{1}{2}mA^2\omega_0^2 = \tfrac{1}{2}kA^2 \tag{6.17}$$

and is constant.

Now, the potential energy at any instant is

$$V = \frac{1}{2}kx^2 = \frac{1}{2}kA^2 \sin^2 [\omega_0(t - t_0) + \theta_0] \tag{6.18}$$

This gives maximum potential energy

$$V_{max} = \frac{1}{2}kA^2$$

and is equal to the maximum kinetic energy.

At any instant, the total energy of the particle is the sum of the kinetic and potential energies at that 'instant'. Thus

$$E = T + V = \frac{1}{2}kA^2 \cos^2 [\omega_0(t - t_0) + \theta_0] + \frac{1}{2}kA^2 \sin^2 [\omega_0(t - t_0) + \theta_0]$$

or

$$E = \frac{1}{2}kA^2 = \frac{1}{2}m\omega_0^2 A^2 \tag{6.19}$$

Thus, the total energy of the harmonic oscillator at any instant is constant and is proportional to the square of the amplitude and to the square of the frequency of the oscillator. The energy changes from kinetic to potential and vice-versa as the oscillator oscillates.

We shall now consider some simple examples of simple harmonic oscillations.

(b) Simple Pendulum

A simple pendulum consists of a heavy particle suspended by a weightless and inextensible rod from a rigid support and is free to oscillate under the action of gravity about an axis perpendicular to the suspension rod. It is an example of a simple harmonic oscillator when the displacement of the pendulum from its equilibrium position is small as compared to its length.

Let OA represent the equilibrium position of the pendulum. Let its displaced position be OB (Fig. 6.1), such that angle AOB is small (about 3° to 5°). The forces acting on the bob are (i) its weight mg acting vertically downwards, and (ii) tension T in the string acting along the string towards O.

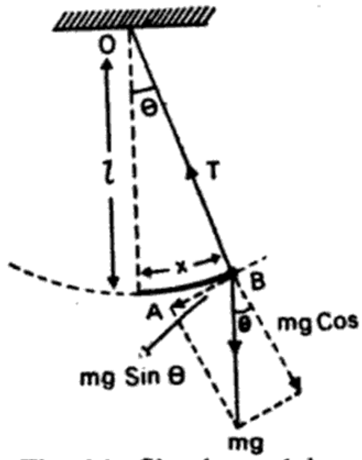


Fig. 6.1 Simple pendulum

Out of the two components of weight mg — $mg \cos \theta$ parallel to the string and $mg \sin \theta$ perpendicular to the string—the perpendicular component $mg \sin \theta$ is responsible for restoring the particle to its original position A. Hence, the equation of motion is

$$m\ddot{x} = -mg \sin \theta \tag{6.20}$$

If θ is small, $\sin \theta \simeq \theta$ (in radians) and equation (6.20) reduces to

$$m\ddot{x} = -g\frac{x}{l} \tag{6.21}$$

where $\theta = \frac{x}{l}$ and x is the displacement along the path of the pendulum which is, in this approximation, nearly straight.

The negative sign on the right-hand side is used to indicate that the force and hence acceleration $\ddot{x}$ are always directed opposite to displacement x which is measured from A.

Equation (6.21) is similar to equation (6.7) and represents an undamped linear simple harmonic motion with the solutions obtained above. Thus, the bob of the simple pendulum performs linear simple harmonic motion with the force constant $k = \frac{mg}{l}$ and, hence, has the period of oscillation given by the well-known formula

$$\tau = 2\pi\sqrt{l/g} \tag{6.22}$$

In obtaining equation (6.21) we assumed that the amplitude of oscillation is small enough to enable us to replace $\sin\theta$ by θ. In this approximation we are really retaining the first term in the expansion

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \ldots \tag{6.23}$$

We now wish to drop this restriction and find the period for relatively larger amplitudes. The equation of motion (6.20) is now written as

$$ml\ddot{\theta} = -mg\sin\theta \tag{6.24}$$

where $\ddot{x} = l\ddot{\theta}$ and $\ddot{\theta}$ is the angular acceleration

or
$$\ddot{\theta} = -\frac{g}{l}\sin\theta \tag{6.25}$$

This is a non-linear, second-order differential equation and cannot be solved exactly in terms of the elementary functions. We shall employ the method of series expansion in evaluating the integrals encountered in finding the solution of equation (6.25).

Multiplying both sides of equation (6.25) by the corresponding sides of the identity $\dot{\theta}\,dt = \frac{d\theta}{dt}\,dt = d\theta$, we get

$$\ddot{\theta}\dot{\theta}\,dt = -\frac{g}{l}\sin\theta\,d\theta,$$

i.e.
$$\tfrac{1}{2}d(\dot{\theta}^2) = \frac{g}{l}\,d(\cos\theta) \tag{6.26}$$

Integrating equation (6.26) and using the fact that the pendulum comes to rest momentarily when its displacement is maximum, i.e., when $\theta = \alpha$, $\dot{\theta} = 0$, we get

$$\dot{\theta}^2 = \frac{2g}{l}(\cos\theta - \cos\alpha) \tag{6.27}$$

or
$$\frac{d\theta}{\sqrt{\cos\theta - \cos\alpha}} = \sqrt{\frac{2g}{l}}\,dt \tag{6.28}$$

Now, the time required by the pendulum to swing from $+\alpha$ to $-\alpha$ is equal to half the period of oscillation τ_α for the amplitude α. Hence

$$\int_{-\alpha}^{+\alpha} \frac{d\theta}{\sqrt{\cos\theta - \cos\alpha}} = \sqrt{\frac{2g}{l}} \int_0^{\tau_\alpha/2} dt = \sqrt{\frac{2g}{l}} \cdot \frac{\tau_\alpha}{2} \qquad (6.29)$$

With the substitution of

$$\cos\theta = 1 - 2\sin^2\frac{\theta}{2}$$

and

$$\cos\alpha = 1 - 2\sin^2\frac{\alpha}{2}$$

in equation (6.28), we get

$$\int_{-\alpha}^{+\alpha} \frac{d\theta}{\sqrt{\sin^2\frac{\alpha}{2} - \sin^2\frac{\theta}{2}}} = \sqrt{\frac{g}{l}}\,\tau_\alpha \qquad (6.30)$$

The left-hand side is an elliptic integral.

To evaluate the left-hand side, we simplify it by substituting

$$\sin\frac{\theta}{2} = \sin\frac{\alpha}{2}\sin\varphi$$

Then, at $\theta = -\alpha$, $\varphi = -\frac{\pi}{2}$ and at $\theta = \alpha$, $\varphi = \frac{\pi}{2}$

Further

$$\tfrac{1}{2}\cos\frac{\theta}{2}\,d\theta = \sin\frac{\alpha}{2}\cos\varphi\,d\varphi$$

Hence

$$d\theta = \frac{2\sin\frac{\alpha}{2}\cos\varphi\,d\varphi}{\cos\frac{\theta}{2}}$$

or

$$d\theta = \frac{2\sin\frac{\alpha}{2}\cos\varphi\,d\varphi}{\sqrt{1 - \sin^2\frac{\alpha}{2}\sin^2\varphi}} \qquad (6.31)$$

With these substitutions, equation (6.30) becomes

$$\begin{aligned}
\sqrt{\frac{g}{l}}\,\tau_\alpha &= \int_{-\pi/2}^{\pi/2} \frac{2d\varphi}{\sqrt{1 - \sin^2\frac{\alpha}{2}\sin^2\varphi}} \\
&= \int_{-\pi/2}^{\pi/2} 2d\varphi \left(1 - \sin^2\frac{\alpha}{2}\sin^2\varphi\right)^{-1/2} \\
&= \int_{-\pi/2}^{\pi/2} 2d\varphi \left(1 + \tfrac{1}{2}\sin^2\frac{\alpha}{2}\sin^2\varphi + \tfrac{3}{8}\sin^4\frac{\alpha}{2}\sin^4\varphi + \ldots\right) \\
&= 2\left[\pi + \tfrac{1}{2}\sin^2\frac{\alpha}{2}\left(\frac{\pi}{2}\right) + \tfrac{3}{8}\sin^4\frac{\alpha}{2}\left(\frac{3\pi}{8}\right) + \ldots\right]
\end{aligned}$$

Hence

$$\tau_\alpha = \tau\left[1 + \tfrac{1}{4}\sin^2\frac{\alpha}{2} + \frac{9}{64}\sin^4\frac{\alpha}{2} + \ldots\right] \qquad (6.32)$$

where, $2\pi\sqrt{\frac{l}{g}} = \tau$, the period of the simple pendulum when the angular amplitude is small.

If $\alpha = 90°$,

$$\tau_{90°} = \tau[1 + \tfrac{1}{8} + \tfrac{9}{256} + \ldots]$$
$$= 1.16\tau \tag{6.33}$$

But, if α is nearly 20°, the term $\sin^4\frac{\alpha}{2}$ becomes small and $\sin^2\frac{\alpha}{2}$ can be replaced by $\frac{\alpha^2}{4}$. With this limitation, we get

$$\tau_\alpha = \tau\left[1 + \frac{\alpha^2}{16}\right]$$

or

$$\tau = \tau_\alpha\left[1 - \frac{\alpha^2}{16}\right] \tag{6.34}$$

It can be shown that for $\alpha = 4°$,

$$\tau_\alpha = 1.0003\tau$$

Thus, the periods differ by 3 parts in 10000.

(c) Angular Simple Harmonic Motion—Compound Pendulum

A rigid body capable of swinging in a vertical plane about an axis passing through any point in it forms a compound pendulum. If the rigid body is displaced from the equilibrium position, a restoring couple acts on it and the body performs angular oscillations. If angle of rotation θ is small, the restoring couple is proportional to θ and the oscillations are simple harmonic. Then, the equation of motion is written as

$$I\ddot{\theta} = -c\theta \tag{6.35}$$

where I is the M.I. of the body about the given axis of rotation and c is the restoring torque per unit angular displacement.

Equation (6.35) can be written as

$$\ddot{\theta} = \frac{-c}{I}\theta$$

The period of this angular simple harmonic motion is given by the formula

$$\tau = 2\pi\sqrt{\frac{I}{c}} \tag{6.36}$$

Any rigid body capable of rotation about an axis can be called a compound or physical pendulum. Let m be the mass of the body. Then, from Fig. 6.2, it is obvious that the restoring torque is $mgl \sin\theta$, where l is the distance of the centre of gravity from the axis of rotation. If θ is small, the restoring torque is $mgl\theta$. Then, c, the restoring torque per unit angular displacement is equal to mgl.

Further, by using the theorem of parallel axes, the moment of inertia

$$I = I_G + ml^2$$
$$= m(K^2 + l^2)$$

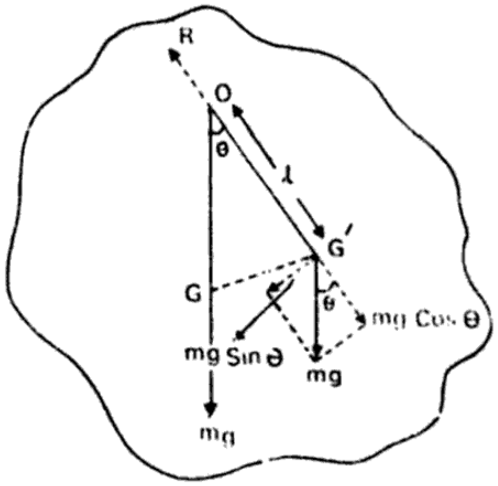


Fig. 6.2 A rigid body oscillating in a vertical plane about O

where K is the radius of gyration about the centre of gravity.

Hence, the period of the compound pendulum becomes

$$\tau = 2\pi\sqrt{(K^2 + l^2)/lg} = 2\pi\sqrt{L/g}$$

where $L = (K^2 + l^2)/l$.

This length L is called the length of an equivalent simple pendulum.

It can be easily shown that a physical pendulum is a reversible pendulum, i.e., we can always find point O' in the pendulum about which it will have the same period of oscillation. A bar pendulum or a Kater's pendulum are examples of compound pendulum.

(d) Oscillations of a Mass Attached to the Spiral Spring

Consider a spiral spring of initial length L. Let it be extended by length l when load mg is attached to it. If the force constant of the spring is k, then the spring is in equilibrium under the forces kl and mg, and so, $k = \frac{mg}{l}$ in magnitude.

If now the mass is pulled down slightly and then released, it performs the oscillations in the vertical direction. Equation (6.7) describes this simple harmonic motion and the period of oscillation is given by

$$\tau_0 = 2\pi\sqrt{\frac{m}{k}} = 2\pi\sqrt{\frac{l}{g}} \tag{6.37}$$

Thus, the period of oscillation of the loaded spiral spring is equal to that of a simple pendulum whose length is equal to the extension produced in the spring.

Effect of the Mass of the Spring: In the above derivation the mass of the spring was neglected in comparison with the mass that was attached to it. We now find the effect of the mass of the spring on the period. For this, we calculate the kinetic energy of the system consisting of the spring and the mass attached to it.

Let us assume that the spring is uniformly wound and that its length is L and mass is M. Then, the mass per unit length σ is $\sigma = M/L$.

Consider a small element dy of the spring at height y above the equilibrium position (Fig. 6.3). The mass of this element is $\sigma\, dy$. Let v be the

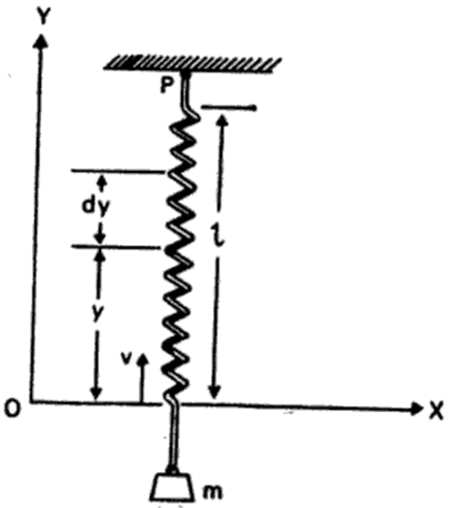


Fig. 6.3 Kinetic energy of spring and mass m

velocity of the lowest point of the spring at any instant. The velocity of the other end of the spring is always zero. Hence, the velocity of the small element under consideration is

$$v_y = \frac{v(L-y)}{L}$$

Therefore, the kinetic energy of the element is

$$\tfrac{1}{2}\sigma\, dy\, v_y^2 = \tfrac{1}{2}\sigma \frac{v^2(L-y)^2}{L^2}\, dy$$

Hence, the total kinetic energy of the mass and the spring is given by

$$T = \tfrac{1}{2}mv^2 + \tfrac{1}{2}\frac{\sigma v^2}{L^2}\int_0^L (L-y)^2\, dy$$

$$= \tfrac{1}{2}mv^2 + \tfrac{1}{2}\frac{M}{3}v^2$$

Thus
$$T = \tfrac{1}{2}\left(m + \frac{M}{3}\right)v^2$$

Thus, one-third the mass of the spring is effectively added to mass m attached at the end of the spring.

The formula for period must then be modified to

$$\tau' = 2\pi\sqrt{\frac{(m + M/3)}{k}} \tag{6.38}$$

to include the effect of mass of the spring.

6.2 DAMPED HARMONIC OSCILLATOR

Suppose that a resistive or a damping force is present in addition to the

restoring force which is necessary to produce the oscillations. We assume that this damping force is proportional to the velocity of the particle and is given by $F_{\text{damping}} = -2m\mu\dot{x}$, $(\mu > 0)$ Quantity $2m\mu$ represents the damping force per unit velocity and is so chosen to simplify the calculations as will be seen presently. The equation of motion, then, becomes

$$m\ddot{x} = -kx - 2m\mu\dot{x} \tag{6.39}$$

or

$$\ddot{x} + 2\mu\dot{x} + \omega_0^2 x = 0 \tag{6.40}$$

where $\omega_0^2 = \dfrac{k}{m}$ as before.

Equation (6.40) is a homogeneous linear differential equation of the second order with constant coefficients. Let us try a solution of the equation in the form

$$x \propto e^{\alpha t}$$

where α is constant, to be determined. Substituting for x, $\dot{x}$ and $\ddot{x}$ in equation (6.40), we get

$$\alpha^2 + 2\mu\alpha + \omega_0^2 = 0 \tag{6.41}$$

an equation quadratic in α and has two roots given by

$$\alpha = -\mu \pm \sqrt{\mu^2 - \omega_0^2} = \mu \pm \lambda \tag{6.42}$$

where $\lambda = \sqrt{\mu^2 - \omega_0^2}$.

Thus, for all $\mu \neq 0$, we have two solutions

$$x_1 = Ae^{-(\mu+\lambda)t}$$

and

$$x_2 = Be^{-(\mu-\lambda)t}$$

The general solution is the linear sum of the two solutions. Hence, we get

$$x = x_1 + x_2 = Ae^{-(\mu+\lambda)t} + Be^{-(\mu-\lambda)t}$$

i.e.

$$x = e^{-\mu t}(Ae^{-\lambda t} + Be^{+\lambda t}) \tag{6.43}$$

as the general solution of equation (6.40).

The solution given by equation (6.43) represents different physical situations depending upon relative values of μ and ω_0. We have the following three cases accordingly.

Case I: *Overdamped motion*: If $\mu > \omega_0$, we find that $\mu > \lambda$ as well.

Hence, both the terms on the right-hand side of equation (6.43) decay exponentially, the first one decreasing faster than the other.

Let us have initial conditions $x = x_0$ and $\dot{x} = v_0$ at $t = 0$.

Then, from equation (6.43) and its time derivative, we get at $t = 0$

$$A + B = x_0$$

and

$$-(\mu + \lambda)A - (\mu - \lambda)B = v_0$$

From these expressions we get

$$A = -\frac{v_0 + (\mu - \lambda)x_0}{2\lambda} \tag{6.44}$$

and

$$B = \frac{v_0 + (\mu + \lambda)x_0}{2\lambda} \tag{6.45}$$

Thus, when x_0 and v_0 are positive, B is +ve and A is −ve. Moreover, the magnitude of B is greater than that of A.

Displacement x, therefore, will be positive at all instants as is qualitatively represented by curve (a) of Fig. 6.4.

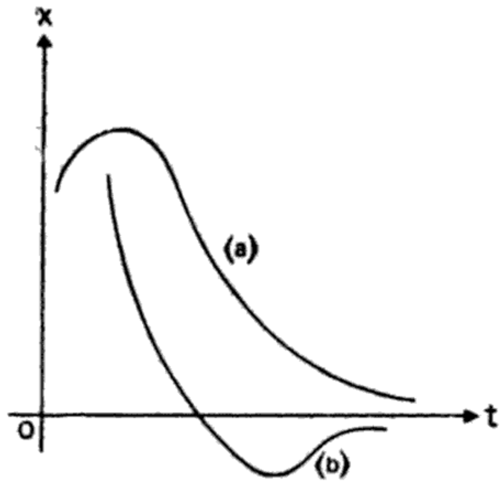


Fig. 6.4 Overdamped motion

If, on the other hand, v_0 is negative such that $v_0 < -(\mu + \lambda)x_0$, B is negative and A is positive. Furthermore, magnitudewise A is greater than B. Since the term containing A decays more rapidly than the term containing B, the term containing B will be predominant after some time when the term containing A becomes insignificant. Thus, the positive displacement will become negative crossing the equilibrium position and once again will tend monotonically to the equilibrium position as is shown by curve (b) in Fig. 6.4.

Case II: *Underdamped motion*: If $\mu < \omega_0$, $\sqrt{\mu^2 - \omega_0^2}$ is an imaginary quantity. Call it $i\omega'$, where $\omega' = \sqrt{\omega_0^2 - \mu^2}$ and $i = \sqrt{-1}$. The general solution given by equation (6.43) then becomes

$$x = e^{-\mu t}(Ae^{i\omega' t} + Be^{-i\omega' t}) \tag{6.46}$$

It should be noted that x is a real quantity, being the displacement. Hence, constants A and B will be complex numbers and must be related in such a way as to make the right-hand side real. To obtain the values of A and B, we write equation (6.46) in the form

$$\begin{aligned} x &= e^{-\mu t}[A(\cos \omega' t + i \sin \omega' t) + B(\cos \omega' t - i \sin \omega' t)] \\ &= e^{-\mu t}[(A + B) \cos \omega' t + i(A - B) \sin \omega' t] \end{aligned} \tag{6.47}$$

Let $A + B = C \sin \varphi$ and $i(A - B) = C \cos \varphi$, where C and φ are constants—both real numbers. Substituting these values in equation (6.47), we get

$$x = Ce^{-\mu t} \sin (\omega' t + \varphi) \tag{6.48}$$

Equation (6.48) shows that the system performs sinusoidal oscillations having a period

$$\tau' = \frac{2\pi}{\omega'} = \frac{2\pi}{\sqrt{\omega_0^2 - \mu^2}}$$

and amplitude $Ce^{-\mu t}$.

Thus, we observe that:

(i) period τ' in the presence of a damping force is different from the natural period $\tau_0 = (2\pi/\omega_0)$, and

(ii) the amplitude of oscillations, viz. $Ce^{-\mu t}$ decreases exponentially with time.

Such oscillations are represented in Fig. 6.5.

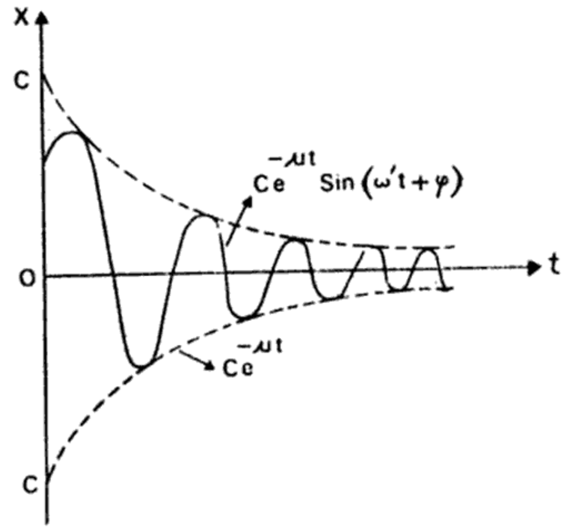


Fig. 6.5 Damped oscillations

The ratio of the amplitude at two successive points separated by time τ' is given by

$$\frac{Ce^{-\mu t}}{Ce^{-\mu(t+\tau')}} = e^{\mu\tau'}$$

Quantity $\mu\tau'$ is called the logarithmic decrement of the motion and $\mu\tau' = \dfrac{2\mu\pi}{\sqrt{\omega_0^2 - \mu^2}}$.

Case III: *Critically damped motion*: If $\mu \simeq \omega_0$, we can get the result directly by substituting this equality or $\lambda = 0$ in equation (6.43). We have to treat this as a limiting case when $\mu \to \omega_0$. For this purpose, let $\sqrt{\mu^2 - \omega_0^2} = h$, where h is a very small quantity.

Then, the general solution given by equation (6.43) becomes

$$\begin{aligned} x &= e^{-\mu t}[A(1 + ht) + B(1 - ht)] \\ &= e^{-\mu t}[(A + B) + (A - B)h] \end{aligned}$$

where we have used

$$e^{ht} = 1 + ht + \ldots$$

and

$$e^{-ht} = 1 - ht + \ldots$$

and neglected the higher-order terms as they are small.

Therefore

$$x = e^{-\mu t}[G + Ht] \tag{6.49}$$

where $G = A + B$ and $H = (A - B)h$

It will be observed that for the same initial conditions, the system will return to the equilibrium position most quickly if it is critically damped. This will not be the case if $B = 0$.

Term $Hte^{-\mu t}$ decays less rapidly than term $Ge^{-\mu t}$. Hence, depending upon the relative magnitudes and signs of G and H, we get cases similar to those obtained in the case of overdamped motion. The corresponding curves also are, therefore, more or less similar.

Energy Dissipation

The loss of energy at any instant due to the damping force per unit time is given by R, where

$$R = \text{damping force} \times \text{velocity} = 2m\mu\dot{x}^2, \text{ numerically.}$$

Now, the equation of motion of the damped system, after multiplying throughout by $\dot{x}$, is

$$m\dot{x}\ddot{x} + 2m\mu\dot{x}^2 + \omega_0^2 x\dot{x} = 0$$

i.e.

$$\begin{aligned} 2m\mu\dot{x}^2 &= -[m\dot{x}\ddot{x} + \omega_0^2 x\dot{x}] \\ &= -\left[\frac{m}{2}\frac{d}{dt}(\dot{x}^2) + \frac{\omega_0}{2}\frac{d}{dt}(x^2)\right] \\ &= -\left[\frac{d}{dt}(\tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}\omega_0 x^2)\right] = -\frac{d}{dt}(T + V) = -\frac{dE}{dt} \end{aligned}$$

where T and V represent the kinetic and the potential energies, respectively, and the total energy $E = T + V$. Hence

$$R = 2m\mu\dot{x}^2 = -\frac{dE}{dt}$$

or

$$\frac{dE}{dt} = -2m\mu\dot{x}^2 \tag{6.50}$$

Thus, the total energy goes on decreasing due to damping and the rate of decrease is proportional to the square of the velocity when the damping force is proportional to velocity.

If the damping force is zero, the right-hand side of equation (6.50) will be zero and we have

$$\frac{dE}{dt} = 0 \quad \text{or} \quad E = T + V = \text{const} \tag{6.51}$$

Equation (6.51) shows that the total energy of an undamped oscillator is conserved and the energy of the damped oscillator decreases according to equation (6.50). The energy lost by the oscillator is due to frictional forces in the medium and goes into heating it.

6.3 FORCED OSCILLATIONS

We saw in the last article that a damped system would oscillate under

This can be seen from the following considerations. When amplitude is half of A_{max}

$$\frac{A_{max}}{2} = A(\text{at } \omega_h = \omega_r \pm \Delta)$$

or from equations (6.58) and (6.62)

$$\frac{F_0/m}{2\sqrt{4\mu^2\omega_0^2 - 4\mu^4}} = \frac{F_0/m}{\sqrt{(\omega_0^2 - \omega_h^2)^2 + 4\mu^2\omega_h^2}}$$

Hence, solving this equation, we find that the half-maximum value of A occurs at a frequency given by

$$\omega_h^2 = \omega_r^2 \pm 2\sqrt{3}\mu\sqrt{(\omega_0^2 - \mu^2)}$$

or for small values of μ, $(\omega_r \approx \omega_0)$

$$\omega_h \simeq \omega_r \pm \sqrt{3}\mu$$

wherein we have neglected the terms of the order of μ^2.

Thus

$$|\omega_h - \omega_r| = \Delta = \mu\sqrt{3} \tag{6.70}$$

This shows that for small μ, ω_h is close to ω_r, Δ is small and the resonance is sharp.

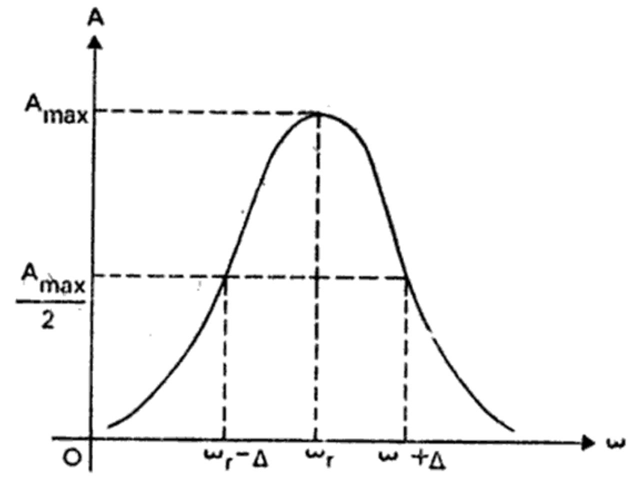


Fig. 6.8 Resonance curve—half-width

The sharpness of resonance is also measured in terms of another quantity called the Q-factor of the system. It is defined as 2π times the ratio of the average energy stored in the system to the energy dissipated per cycle by the applied force.

Hence

$$Q = \frac{2\pi\langle\frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega_0^2x^2\rangle}{\tau\langle 2m\mu\dot{x}^2\rangle}$$

for a damped simple harmonic oscillator.

Here, we have used the notation $\langle\,\rangle$ to denote time averge of the quantity. The average value of $\dot{x}^2$ is

$$\langle\dot{x}^2\rangle = \langle A^2\omega^2\cos^2(\omega t - \theta)\rangle$$
$$= \tfrac{1}{2}A^2\omega^2$$

Similarly,

$$\langle x^2 \rangle = \frac{A^2}{2}$$

where we have used the fact that the time average of $\sin^2(\omega t - \theta)$ and $\cos^2(\omega t - \theta)$ over a cycle is $\frac{1}{2}$. Hence, the Q-factor is given by

$$Q = \frac{\omega_0^2 + \omega^2}{4\mu\omega} \tag{6.71}$$

where we used $\tau = \dfrac{2\pi}{\omega}$

Near resonance $\qquad \omega \simeq \omega_0$ and $Q = \dfrac{\omega_0}{2\mu}$

This shows that the value of Q is greater for smaller values of μ and corresponds to greater sharpness of the resonance. Hence, we conclude that the greater the value of Q, the greater is the sharpness of resonance.

(d) Phase Relationships

We have seen that if force $F = F_0 \sin \omega t$ is applied to a damped harmonic oscillator, the system oscillates and the displacement and velocity are given by

$$x = A \sin(\omega t - \theta)$$

and

$$\dot{x} = A\omega \cos(\omega t - \theta)$$

or

$$\dot{x} = A\omega \sin(\omega t - \gamma)$$

where $\gamma = \theta - \dfrac{\pi}{2}$ and represents phase difference between the applied force and velocity.

Hence from equation (6.57), the phase difference between the applied force and the velocity is given by γ where

$$\tan \gamma = -\cot \theta = \frac{\omega^2 - \omega_0^2}{2\mu\omega} \tag{6.72}$$

For velocity resonance

$$\omega^2 = \omega_0^2 = \frac{k}{m}$$

Hence, $\qquad \tan \gamma = 0 \quad \text{or} \quad \gamma = 0$

Thus, at the velocity resonance, the velocity and the applied force are in phase whereas the displacement lags behind the applied force by 90°.

Velocity resonance is of great importance in alternating current circuits, the current being analogous to the velocity.

(e) Energy Considerations in Forced Oscillations

In the case of forced oscillations, the driving force supplies energy to the system and maintains the oscillations. Hence, in the steady state, when oscillations are maintained at constant amplitude, it is expected that the driving force supplies the energy to the system at a rate equal to that at which it is dissipated in doing work against the damping force. We shall prove this statement.

The instantaneous rate of supply of energy

$$= \text{velocity} \times \text{driving force}$$
$$= A\omega \sin(\omega t - \gamma) \times F_0 \sin \omega t$$
$$= A\omega F_0 \sin^2 \omega t \cos \gamma - \tfrac{1}{2} A\omega F_0 \sin 2\omega t \sin \gamma$$

The time average value of $\sin^2 \omega t$ over a complete cycle is $\frac{1}{2}$, while that of $\sin 2\omega t$ is zero.

Hence, the time average rate at which energy is supplied to the system is $\frac{1}{2} A\omega F_0 \cos \gamma$. Call it $\langle \dot{W} \rangle$.

Then $$\langle \dot{W} \rangle = \frac{A\omega F_0 \cos \text{[illegible]}}{2}$$

But $$\cos \gamma = \cos\left(\theta - \frac{\pi}{2}\right) = \sin \theta = \frac{\text{[illegible]}}{\sqrt{(\omega_0^2 - \text{[illegible]}}}$$

from equation (6.57), and amplitude is

$$A = \frac{F_0/m}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\mu^2\omega^2}}$$

Hence, A can be expressed as

$$A = (F_0/m)(\cos \gamma / 2\mu\omega)$$

Hence $$\langle \dot{W} \rangle = \frac{F_0^2 \cos^2 \gamma}{4\mu m} \tag{6.73}$$

Similarly, the rate of dissipation of energy is given [illegible]

$$D = \text{velocity} \times \text{damping force}$$
$$= 2\mu m \dot{x}^2$$
$$= 2\mu m A^2 \omega^2 \sin^2(\omega t - \gamma)$$

Hence, the time average rate of dissipation of energy is given by

$$\langle D \rangle = 2\mu m A^2 \omega^2 \times \tfrac{1}{2}$$
$$= \mu m \frac{F_0^2}{m^2[(\omega_0^2 - \omega^2)^2 + 4\mu^2\omega^2]} \omega^2$$
$$= \frac{\mu F_0^2}{m} \frac{\cos^2 \gamma}{4\mu^2\omega^2} \omega^2$$

or $$D = \frac{F_0^2 \cos^2 \gamma}{4\mu m} \tag{6.74}$$

Equations (6.73) and (6.74) prove the statement.

6.4 COUPLED OSCILLATIONS

We now discuss the motion of two oscillating systems in which the motion of one system can influence that of the other. The two systems are then said to be coupled to each other and the resulting oscillations are called *coupled oscillations*. The extent to which the motion of one system influences the motion of the other is called the 'coupling' of the systems. The coupled systems may be mechanical or electrical. We restrict our attention to the systems, the coupling between which is weak.

Consider a system of two identical simple pendulums coupled by means of a spring (Fig. 6.9). In such a system, there is resonance between the

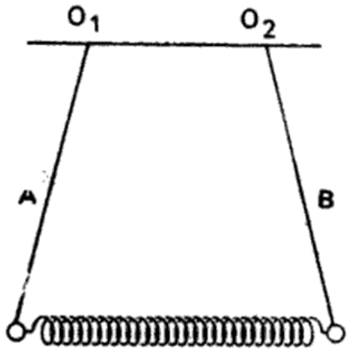


Fig. 6.9 Coupled system

two parts of the system. Suppose that pendulum A oscillates first with pendulum B at rest. The energy from A gets transferred to B and it alternates between the pendulums A and B. When pendulum A has maximum amplitude, pendulum B has zero amplitude and vice versa (Fig. 6.10).

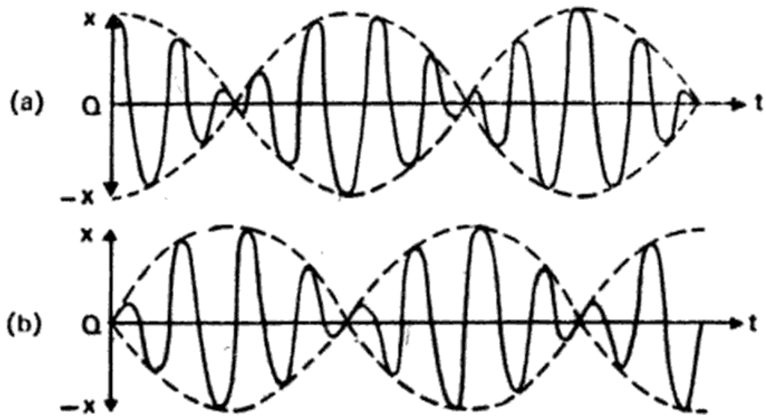


Fig. 6.10 (a) Motion of pendulum A, (b) Motion of pendulum B

Consider two identical simple pendulums each having length l and a bob of mass m coupled to each other by means of a horizontal spring (Fig. 6.11). The mass of the spring is assumed to be negligible. Let the force constant of the spring be k.

Let the spring be unstretched when the two bobs are in the equilibrium position.

At some later instant t, let x_1 and x_2 represent the displacements of the two bobs from their equilibrium positions. Then, the spring is stretched by amount $(x_2 - x_1)$ and supplies force $k(x_2 - x_1)$. Let us assume that the motion is undamped and the displacements are small. Then, the restoring force due to mg will be written as $\frac{mgx}{l}$ as in the case of a simple pendulum.

differentiating the solution with respect to time. Thus

$$\dot{x} = \frac{(F_0/m)\omega \cos(\omega t - \theta)}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\mu^2\omega^2}} = A\omega \cos(\omega t - \theta) \tag{6.63}$$

where we have put

$$A\omega = \frac{F_0/m}{\sqrt{\left(\frac{\omega_0^2 - \omega^2}{\omega}\right)^2 + 4\mu^2}} \tag{6.64}$$

The velocity amplitude will be maximum, when the denominator is minimum. The only quantity in the denominator that can be changed is $\left(\frac{\omega_0^2 - \omega^2}{\omega}\right)^2$, since μ = constant and ω can be changed by changing frequency of the force. But, it is a squared quantity and hence its minimum value must be zero. This gives

$$\omega = \omega_0 \tag{6.65}$$

or

$$\nu = \nu_0 \tag{6.66}$$

Thus, velocity resonance occurs when the inducing frequency equals the natural frequency of the oscillator. It should be noted that this frequency and the frequency of amplitude resonance given by equation (6.60) or (6.61) are not the same in general. The two frequencies are equal for free oscillations.

Further, the maximum value of $A\omega$ will be given by

$$(A\omega)_{\max} = \frac{F_0/m}{2\mu} = \frac{F_0}{2m\mu} \tag{6.67}$$

Thus, $(A\omega)_{\max}$ also depends upon the damping factor and becomes infinite for free oscillations ($\mu = 0$). The variation of $A\omega$ with ω for various values of μ is qualitatively similar to the curves shown in Fig. 6.7, for amplitude resonance.

(c) Sharpness of Resonance

We have seen above that the displacement resonance or the velocity resonance takes place at a particular frequency. The resonance curves drawn in Fig. 6.7 show that the smaller the value of the damping factor, the greater is amplitude $A_{\max}$ or $(A\omega)_{\max}$ and the steeper is the resonance curve. Thus, the steepness or sharpness of the resonance depends entirely upon the magnitude of the damping factor. We can define a measure for the sharpness as follows:

Consider a resonance curve for the displacement resonance (Fig. 6.8). The amplitude is $A_{\max}$ at frequency ω_r and becomes $\frac{A_{\max}}{2}$ at frequency

$$\omega_h = \omega_r \pm \Delta \tag{6.68}$$

where quantity

$$\Delta = |\omega_h - \omega_r| \tag{6.69}$$

is called the *half-width* of the resonance curve. It can be used to measure the sharpness of resonance. Thus, if Δ is small, the resonance is sharp.

This is called resonance. The frequency at the time of resonance, i.e., the resonant frequency will obviously correspond to the maximum value of amplitude A. Thus, A is maximum when $\frac{dA}{d\omega} = 0$ and $\frac{d^2A}{d\omega^2} < 0$. This condition leads to

$$2(\omega_0^2 - \omega^2) = 4\mu^2$$

or

$$\omega^2 = \omega_0^2 - 2\mu^2$$

with the condition $\omega_0^2 > 2\mu^2$.

Hence, the resonant frequency is

$$\omega_r = \sqrt{\omega_0^2 - 2\mu^2} \tag{6.60}$$

or

$$\nu_r = \frac{\omega_r}{2\pi} = \frac{\sqrt{\omega_0^2 - 2\mu^2}}{2\pi} \tag{6.61}$$

The maximum value of the amplitude is obtained by putting $\omega = \omega_r$ in equation (6.58). Thus

$$A_{\max} = \frac{F_0/m}{\sqrt{4\mu^4 + 4\mu^2(\omega_0^2 - 2\mu^2)}}$$

$$= \frac{F_0/m}{2\mu\sqrt{\omega_0^2 - \mu^2}} \tag{6.62}$$

This shows that the value of $A_{\max}$ depends upon μ, the damping factor. The variation of A with ω for various values of μ as given by equation (6.58) is represented by the curves of Fig. 6.7. Thus, for undamped oscillations ($\mu = 0$) when forcing frequency $\omega = \omega_0$, the natural frequency, the amplitude of the oscillations becomes very large, theoretically infinite.

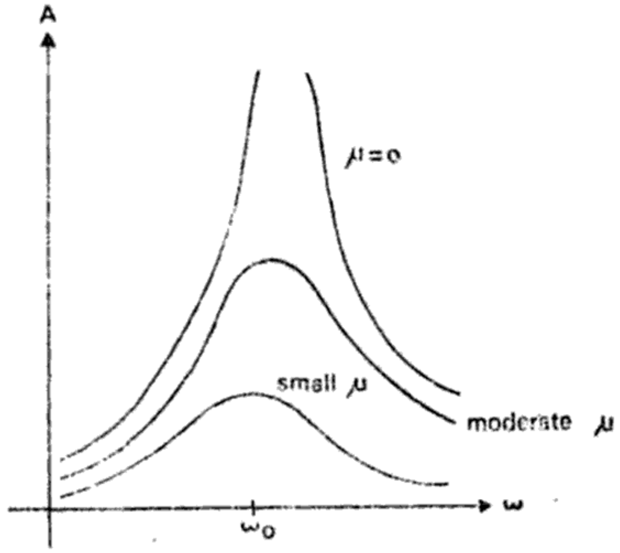


Fig. 6.7 Amplitude resonance

However, due to damping, which is always present in any naturally oscil lating system, the amplitude becomes large but finite.

(b) Velocity Resonance

The velocity of the particle performing forced oscillations is obtained b

This solution of damped forced oscillator shows that the particle performs forced oscillations having amplitude and phase dependent on damping and forcing frequency ω.

Equation (6.59) gives that part of the solution which has the same periodicity as that of the applied force. This corresponds to the inhomogeneous or particular solution. The other solution which corresponds to the homogeneous equation is obtained in the previous article by solving $\ddot{x} + 2\mu\dot{x} + \omega_0^2 x = 0$. But, this leads to oscillations that are damped. This is many times called the transient solution. The particular solution does not die out since oscillations are sustained by the external oscillating force. This solution is, therefore, called the steady state solution. In forced oscillations we are mainly interested in the oscillations given by equation (6.59). From equation (6.57) it is clear that if angular frequency ω of the applied force is increased from zero to natural frequency ω_0 of the system, the phase angle between applied force and displacement increases from zero to 90°. If $\omega > \omega_0$, $\tan\theta$ is negative. Hence, θ takes values greater than 90° and approaches 180° as ω becomes very large as compared to ω_0. In other words, when the frequency of the applied force is very large, the displacement and the applied force are in opposition.

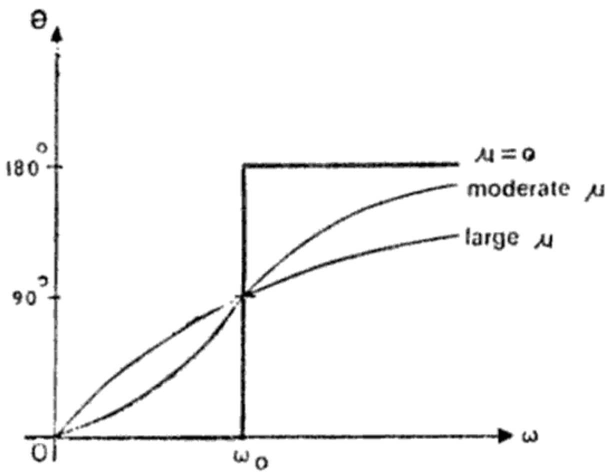


Fig. 6.6 Phase difference between displacement and applied force versus frequency

The variation in phase difference θ with frequency of the applied force is also influenced by the value of μ. This change is shown in Fig. 6.6. If there exists no damping, angle θ changes abruptly from 0 to 180° when frequency ω of the applied force becomes equal to the natural frequency of the system.

(a) Displacement (Amplitude) Resonance

Amplitude A of the forced oscillations varies with the frequency of the applied force. There exists some frequency called the resonance frequency at which the amplitude becomes maximum. At this stage maximum transfer of energy occurs from the driving force to the driven system.

image
not
available

The equations of motion of the two pendulums are written as

$$m\ddot{x}_1 = -\frac{mgx_1}{l} + k(x_2 - x_1) \tag{6.75}$$

and

$$m\ddot{x}_2 = -\frac{mgx_2}{l} - k(x_2 - x_1) \tag{6.76}$$

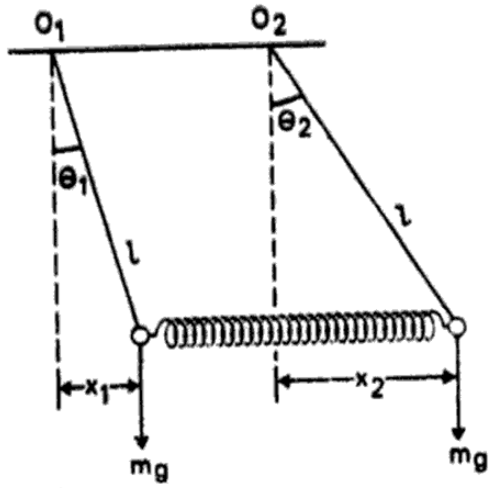


Fig. 6.11 Coordinates of and forces on two pendulums of the coupled system

The elastic force $k(x_2 - x_1)$ due to stretching, pulls the first bob away and the other towards the equilibrium position, hence the elastic force has opposite signs in the above two equations. Let us assume the solutions

$$\left.\begin{aligned} x_1 &= P_1 \cos \omega t + Q_1 \sin \omega t \\ \text{and} \quad x_2 &= P_2 \cos \omega t + Q_2 \sin \omega t \end{aligned}\right\} \tag{6.77}$$

where we have assumed ω to be the angular frequency of each of the two systems. Substituting these values of x_1 and x_2 in equations (6.75) and (6.76) respectively, we get

$$\begin{aligned} &-mP_1\omega^2 \cos \omega t - mQ_1\omega^2 \sin \omega t \\ &= \frac{-mgP_1 \cos \omega t}{l} - \frac{mgQ_1 \sin \omega t}{l} + kP_2 \cos \omega t + kQ_2 \sin \omega t \\ &\qquad - kP_1 \cos \omega t - kQ_1 \sin \omega t \end{aligned} \tag{6.78}$$

and

$$\begin{aligned} &-mP_2\omega^2 \cos \omega t - mQ_2\omega^2 \sin \omega t \\ &= -\frac{mgP_2 \cos \omega t}{l} - \frac{mgQ_2 \sin \omega t}{l} - kP_2 \cos \omega t - kQ_2 \sin \omega t \\ &\qquad + kP_1 \cos \omega t + kQ_1 \sin \omega t \end{aligned} \tag{6.79}$$

The equations (6.78) and (6.79) must hold at all instants. Hence, the coefficients of sine and cosine on the two sides of these equations must be equal. Thus, we get

$$-mP_1\omega^2 = \frac{-mgP_1}{l} + kP_2 - kP_1 \tag{6.80}$$

$$-mQ_1\omega^2 = \frac{-mgQ_1}{l} + kQ_2 - kQ_1 \tag{6.81}$$

and
$$-mP_2\omega^2 = \frac{-mgP_2}{l} - kP_2 + kP_1 \tag{6.82}$$
$$-mQ_2\omega^2 = \frac{-mgQ_2}{l} - kQ_2 + kQ_1 \tag{6.83}$$

The simultaneous equations for P_1 and P_2 can be written as
$$\left(\omega^2 - \frac{g}{l} - \frac{k}{m}\right)P_1 + \frac{k}{m}P_2 = 0 \tag{6.84}$$
$$\frac{k}{m}P_1 + \left(\omega^2 - \frac{g}{l} - \frac{k}{m}\right)P_2 = 0 \tag{6.85}$$

Similar equations hold for Q_1 and Q_2. These two simultaneous equations have non-trivial solution if—
$$\begin{vmatrix} \omega^2 - \frac{g}{l} - \frac{k}{m} & \frac{k}{m} \\ \frac{k}{m} & \omega^2 - \frac{g}{l} - \frac{k}{m} \end{vmatrix} = 0 \tag{6.86}$$

This gives
$$\omega^2 - \frac{g}{l} - \frac{k}{m} = \pm\frac{k}{m}$$

or
$$\omega^2 = \frac{g}{l} + \frac{k}{m} \pm \frac{k}{m} \tag{6.87}$$

Let ω_1 and ω_2 be the two values of ω, then
$$\omega_1^2 = \frac{g}{l} + \frac{2k}{m} \quad \text{or} \quad \omega_1 = \sqrt{\frac{g}{l} + \frac{2k}{m}} \tag{6.88}$$

and
$$\omega_2^2 = \frac{g}{l} \quad \text{or} \quad \omega_2 = \sqrt{\frac{g}{l}} \tag{6.89}$$

Substituting the value of ω_1^2 in equation (6.85), we get $P_2 = -P_1$ and similarly $Q_2 = -Q_1$.

If we use the value of ω_2^2, we would get $P_2 = P_1$ and $Q_2 = Q_1$.

The general solutions can now be written down as
$$\begin{aligned} x_1 &= P_1 \cos \omega t + Q_1 \sin \omega t \\ &= A \cos \omega_1 t + B \sin \omega_1 t + C \cos \omega_2 t + D \sin \omega_2 t \end{aligned} \tag{6.90}$$

and
$$\begin{aligned} x_2 &= P_2 \cos \omega t + Q_2 \sin \omega t \\ &= -A \cos \omega_1 t - B \sin \omega_1 t + C \cos \omega_2 t + D \sin \omega_2 t \end{aligned} \tag{6.91}$$

where A, B, C, and D are the constants of integration and could be found from the initial positions and velocities of the two pendulums.

Each solution here depends on both the frequencies ω_1 and ω_2.

(a) Normal Coordinates

In a coupled system, it is possible to set the two pendulums in motion in such a manner that only one frequency is involved. This can be seen from the following procedure. From equations (6.90) and (6.91), we get
$$x_1 + x_2 = X_1 = 2C \cos \omega_2 t + 2D \cos \omega_2 t \tag{6.92}$$

and
$$x_1 - x_2 = X_2 = 2A \cos \omega_1 t + 2B \cos \omega_1 t \tag{6.93}$$

The modes of oscillation given by the above equations are called the *normal modes* of oscillations for the reason that X_1 and X_2 involve only one frequency ω_2 or ω_1. This is in contrast to the general solutions for x_1 and x_2 which involve both the frequencies ω_2 and ω_1.

Suppose that initially $X_2 = 0$ and the motion represented by X_1 is present (Fig. 6.12a). If the two pendulums are identical, X_2 will remain zero at all instants. Then, $x_1 = x_2$ and there is no oscillation with angular frequency ω_1. There is oscillation with angular frequency ω_2 alone, where $\omega_2 = \sqrt{g/l}$. Hence, the two pendulums oscillate as if they are free. The coupling spring does not get compressed or elongated at any time. This normal mode of oscillation is called a symmetric mode and the particles are oscillating always in phase.

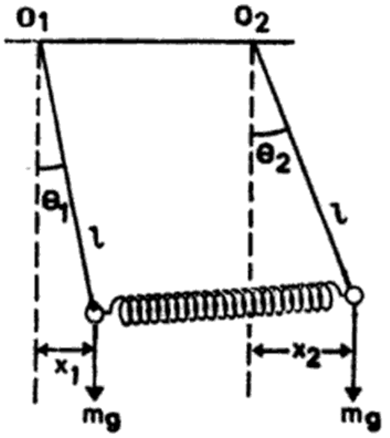


Fig. 6.12a Normal vibrations with ω_2 present and $\omega_1 = 0$

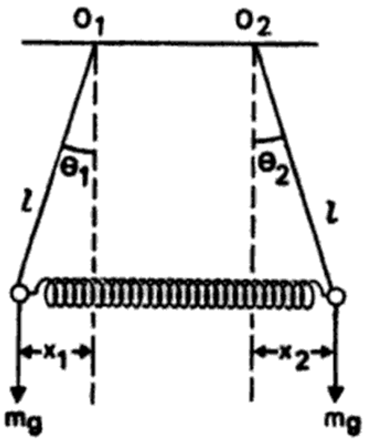


Fig. 6.12b Normal vibrations with ω_1 present and $\omega_2 = 0$

Suppose now that initially, $X_1 = 0$, i.e., $x_1 = -x_2$ (Fig. 6.12b). Then, motion of angular frequency ω_1 alone is present. But

$$\omega_1 = \sqrt{\frac{g}{l} + \frac{2k}{m}}$$

Thus, the spring is compressed and elongated, and exerts a restoring force.

These oscillations are called the antisymmetric mode of oscillations since the particles always move out of phase with frequency ω_1. It should be noted that the antisymmetric mode has a higher frequency of oscillation than that of the symmetric one. In fact this is a very general result and if any complex system with more coupled oscillators is capable of oscillating with different modes with different symmetries, then the mode of oscillations possessing the highest symmetry will have the lowest frequency. Out of the two modes for the two coupled pendulums, the symmetric mode is obviously the one with lower frequency since the spring does not have to do work during the oscillations. When the symmetry is destroyed as in the second case with frequency ω_2, the spring has to do work and hence the frequency is raised.

From equations (6.92) and (6.93), it will be seen that the normal

coordinates X_1 and X_2 are independent of each other, since the system can oscillate with one frequency only while the other is suppressed. In terms of X_1 and X_2, we have

$$x_1 = \frac{X_1 + X_2}{2} \quad \text{and} \quad x_2 = \frac{X_1 - X_2}{2}$$

Substituting these values in equations (6.75) and (6.76), we get

$$\frac{m}{2}(\ddot{X}_1 + \ddot{X}_2) = \frac{-mg}{2l}(X_1 + X_2) - kX_2 \tag{6.94}$$

and

$$\frac{m}{2}(\ddot{X}_1 - \ddot{X}_2) = \frac{-mg}{2l}(X_1 - X_2) + kX_2. \tag{6.95}$$

Adding and subtracting these equations, we get

$$m\ddot{X}_1 = \frac{-mg}{l}X_1 \quad \text{or} \quad \ddot{X}_1 + \frac{g}{l}X_1 = 0 \tag{6.96}$$

and

$$m\ddot{X}_2 = \frac{-mg}{l}X_2 - 2kX_2 \quad \text{or} \quad \ddot{X}_2 + \left(\frac{g}{l} + \frac{2k}{m}\right)X_2 = 0 \tag{6.97}$$

Equations (6.96) and (6.97) represent simple harmonic oscillations with frequencies $\omega_2 = \sqrt{g/l}$ and $\omega_1 = \sqrt{(g/l) + (2k/m)}$ is the normal coordinates, respectively.

(b) Energy of Coupled Oscillations

The total energy of the coupled pendulum system is partly kinetic and partly potential. Further, the potential energy is due to (i) extension of the spring, and (ii) the raising of the pendulum bobs against gravity.

The potential energy due to the extension of the spring is

$$\int_0^{x_2 - x_1} kx\,dx = \frac{k}{2}(x_2 - x_1)^2$$

The potential energy of the pendulum bobs is

$$\int_0^{x_1} \frac{mgx}{l}\,dx + \int_0^{x_2} \frac{mgx\,dx}{l} = \frac{mg}{2l}(x_1^2 + x_2^2)$$

Hence, the total potential energy is

$$V = \frac{k}{2}(x_2 - x_1)^2 + \frac{mg}{2l}(x_1^2 + x_2^2)$$

The kinetic energy of the system is given by

$$T = \tfrac{1}{2}m\dot{x}_1^2 + \tfrac{1}{2}m\dot{x}_2^2$$
$$= \tfrac{1}{2}m(\dot{x}_1^2 + \dot{x}_2^2)$$

The kinetic and potential energies can be expressed in terms of normal coordinates by using $X_1 = x_1 + x_2$ and $X_2 = x_1 - x_2$. Then, we get

$$V = \frac{k}{2}X_2^2 + \frac{mg}{2l}\left[\left(\frac{X_1 + X_2}{2}\right)^2 + \left(\frac{X_1 - X_2}{2}\right)^2\right]$$

or

$$V = \frac{mg}{4l}X_1^2 + \left(\frac{mg}{4l} + \frac{k}{2}\right)X_2^2 \tag{6.98}$$

and

$$T = \tfrac{1}{4}m(\dot{X}_1^2 + \dot{X}_2^2) \tag{6.99}$$

The mechanism of the transfer of energy from one pendulum to the other can be explained as follows. Suppose that pendulum A is initially at rest and is at distance 'a' from the equilibrium position, and pendulum B is at rest in its equilibrium position. Thus, we have

$$x_1 = a, \quad \dot{x}_1 = 0$$

and
$$x_2 = 0, \quad \dot{x}_2 = 0$$

Substituting these initial conditions in equations (6.90) and (6.91), we get

$$a = A + C, \quad 0 = B\omega_1 + D\omega_2$$

and
$$0 = -A + C, \quad 0 = -B\omega_1 + D\omega_2$$

From these we get, $A = C = \frac{a}{2}$ and $B = D = 0$

Hence, equations (6.90) and (6.91) become

$$x_1 = \frac{a}{2}[\cos \omega_1 t + \cos \omega_2 t]$$

and
$$x_2 = \frac{a}{2}[\cos \omega_2 t - \cos \omega_1 t]$$

These can be written as

$$x_1 = a \cos\left(\frac{\omega_1 + \omega_2}{2}\right)t \cos\left(\frac{\omega_1 - \omega_2}{2}\right)t \tag{6.100}$$

and
$$x_2 = a \sin\left(\frac{\omega_1 + \omega_2}{2}\right)t \sin\left(\frac{\omega_1 - \omega_2}{2}\right)t \tag{6.101}$$

If the coupling is weak, ω_1 and ω_2 will not be very much different from each other. Let

$$\frac{\omega_1 + \omega_2}{2} = \omega'$$

Then
$$x_1 = a \cos\left(\frac{\omega_1 - \omega_2}{2}\right)t \cos \omega' t \tag{6.102}$$

and
$$x_2 = a \sin\left(\frac{\omega_1 - \omega_2}{2}\right)t \sin \omega' t \tag{6.103}$$

Equations (6.102) and (6.103) show that x_1 and x_2 are oscillating with angular frequency ω'. But the amplitudes of x_1 and x_2 are slowly varying. The variation is, however, according to a sine factor for x_2 and a cosine factor for x_1. But, we know that as the angle increases, the sine of the angle increases while its cosine decreases. Thus, the amplitudes and hence the energies change in such a manner that the energy of the first pendulum decreases while that of the other increases as angle $\left(\frac{\omega_1 - \omega_2}{2}\right)t$ increases. The energy transfer is periodic and has a period given by

$$\tau = \frac{2\pi}{\left(\frac{\omega_1 - \omega_2}{2}\right)} = \frac{4\pi}{\omega_1 - \omega_2} \tag{6.104}$$

We shall treat the problem of small coupled oscillations by a more general method in Chapter 13.

QUESTIONS

1. Explain the terms: restoring force and force constant.
2. What is meant by forced oscillations? Give some examples of these.
3. Explain the term 'resonance'. How is the resonant frequency affected by damping?
4. What do you understand by 'sharpness of resonance'? Explain how the Q-factor and the half-width are measures of sharpness of resonance.
5. How does the amplitude of oscillation vary around the resonant frequency?
6. A spring having force constant k has mass m suspended from its free end. The spring is cut into two halves and the same mass m is suspended from it. Find the new time period and compare it with the original period.
7. Can we really construct a simple pendulum? Explain.
8. How can we use a pendulum so as to trace out a sinusoidal curve?
9. What component simple harmonic motions would give a figure 8 as the resultant motion?
10. Why are damping devices used on machinery?
11. Give illustrations of resonance.
12. A hollow sphere is filled with water and used as a pendulum bob. If water is allowed to flow slowly through a hole at the bottom, how will the period change?
13. Under what condition is the energy of a simple harmonic oscillator conserved?
14. What is a compound pendulum? Explain the term: length of the equivalent simple pendulum.
15. What is meant by normal coordinates? Explain the terms: symmetric and antisymmetric modes of oscillation.

PROBLEMS

1. A particle moves on a circle of radius r with uniform speed v. Show that the projection of this motion on any diameter is a simple harmonic motion with period $2\pi r/v$.
2. When a particle is subjected to two simple harmonic motions at right angles to each other, say along the x- and y-axes, it produces Lissajous' figures. Suppose that frequency along the y-axis is twice that along the x-axis, the two simple harmonic motions have equal amplitude r and a phase difference of 270°. If they are started with

$x = r$ and $y = 0$, show that the equation of the path is

$$y = \frac{4x^2}{r^2}(r^2 - x^2)$$

3. A particle of mass M is suspended from a spring which has mass M and force constant k. Show that the oscillation frequency is $\omega = \sqrt{3}\omega_0/2$, where ω_0 is the frequency of oscillation if the spring has negligible mass.
4. If the amplitude of a damped oscillator decreases to $1/e$ of its initial value after N periods, show that the frequency of the oscillator is approximately $[1 - (8\pi^2N^2)^{-1}]$ times the frequency of the corresponding undamped oscillator.
5. If a driven (forced) oscillator is only lightly damped, show that the Q-factor of the system is approximately

$$Q \simeq \frac{\text{Total energy}}{\text{Energy loss during one period}}$$

6. For a lightly damped oscillator, show that

$$Q \simeq \omega_0/\Delta\omega$$

where $\Delta\omega$ is the width of the response curve at points for which the amplitude is $\frac{1}{\sqrt{2}}$ of the maximum amplitude.
7. Discuss the motion of a three-dimensional harmonic oscillator which is confined to move in a 'box' having dimensions $2a$, $2b$ and $2c$.
8. An oscillator moves under the influence of a potential

$$V = \tfrac{1}{2}kx^2$$

where k is constant.
Find the period of motion as a function of amplitude.
9. The string of a simple pendulum passes through a small ring, with only the part below the ring free to swing. The length of the pendulum is reduced slowly by pulling the string up. Show that (a) the ratio E/frequency is a constant and is equal to $\oint p\,dx$ and (b) increase in energy is equal to the work done.
10. The natural length of a spring is 10 cm. When a mass is suspended from the end of the spring, the equilibrium length becomes 12 cm. The mass is now given a blow so that it starts moving with velocity 4 cm/s. Find the period and amplitude of the resulting oscillations.
11. A particle moving under a conservative force oscillates between x_1 and x_2. Show that the period of oscillation is

$$\tau = \int_{x_1}^{x_2} \left[\frac{m}{2\{V(x_2) - V(x_1)\}}\right]^{1/2} dx$$

12. A critically damped oscillator of natural period τ is subjected to a periodic force

$$F(t) = c(t - n\tau),\ (n - \tfrac{1}{2})\tau < t < (n + \tfrac{1}{2})\tau$$

where n is an integer and c is a constant. Find the ratios of amplitudes of oscillation at angular frequencies $\frac{2\pi n}{\tau}$.

13. A particle of mass m is suspended at the end of two springs of force constants k_1 and k_2 which are joined together. Find the expression for the frequency.
14. In problem 13, what will be the frequency of oscillation if mass m is supported by both the springs?
15. A particle of mass m is moving in a vertical plane under the action of gravity along the curve
    (i) $y = -l \cos x$
    (ii) $y = -b\sqrt{1 - x^2/a^2}$
    Find the period of small oscillations about the lowest points of the curves.
16. A particle of mass m is attached to a light wire which is stretched tightly between two fixed points. The tension in the string is F. If l_1 and l_2 are the distances of the particle from the two ends, prove that the period of small transverse oscillations is
$$2\pi\sqrt{\frac{ml_1l_2}{(l_1 + l_2)F}}$$
17. Two particles of masses m_1 and m_2 are attached to the ends of a light spring. The natural length of the spring is l and its tension is k times its extension. Initially the particles are at rest with m_1 at height h above m_2. At $t = 0$, m_1 is projected vertically upward with velocity v. Find the positions of the particles at any subsequent time. What is the largest value of v for which this solution applies?
18. Force $F_0(1 - e^{-at})$ acts on a harmonic oscillator which is at rest at $t = 0$. The mass of the oscillator is m, the force constant is $k = 4ma^2$ and the damping force is given by mav. Discuss the motion.
19. An undamped harmonic oscillator of mass m and natural frequency ω_0 is initially at rest. At $t = 0$ it is given a blow so that it starts from $x_0 = 0$ with initial velocity v_0 and oscillates freely until $t = 2\pi/2\omega_0$. From this time on force $f = B \cos(\omega t + \theta)$ is applied. Find the motion.

7
Collisions of Particles

In modern physics, the collision or scattering experiments constitute an important source of information regarding the nature of the interaction between the particles. For example, Rutherford's experiment on scattering of α-particles by atoms in a thin foil of gold revealed the existence of positively charged nucleus in the atom.

When two particles approach each other, a force of interaction comes into play and the particles get scattered. Even if the nature of this force of interaction is not known, we can apply the laws of conservation of linear momentum and energy and write down the relations between the initial and the final momenta and the energies of the particles. In an experimental set up, we know the initial conditions, i.e., the conditions of the particles when they are so far away from each other that the interaction is yet to start. One of the particles is usually at rest and the other approaches it. The interaction between the particles takes place in a very short interval of time. During this interval, forces of interaction come into play and the trajectory of the particle is changed. In the final condition, the interacting particles are once again far away from each other and can be treated as free particles. The angle through which the incident particle is deflected is called the angle of scattering. It depends upon the nature of the interaction and its value cannot be predicted. However, it can easily be determined experimentally.

The force of interaction may be due to different causes in different cases. Thus, in the collision between two billiard balls, the force of interaction is due to elasticity. It comes into existence only when the two billiard balls come into physical contact. In the case of scattering of alpha particles by the nuclei, it is the electrostatic force that causes the interaction. It is the gravitational force that causes the two stars to deflect each other and so on.

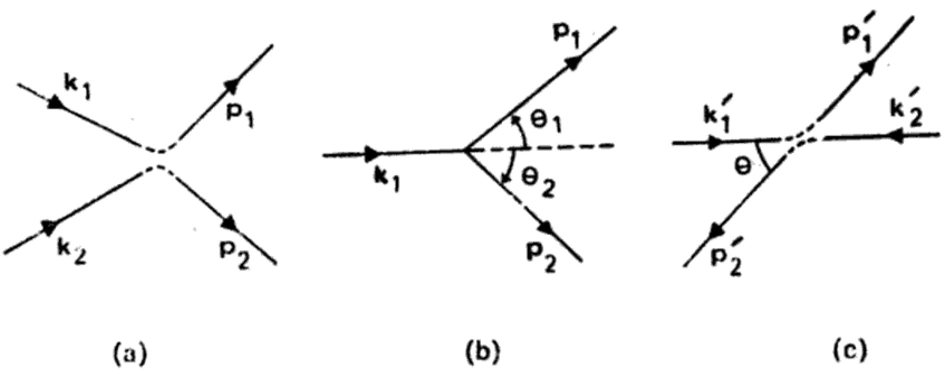
7.1 ELASTIC AND INELASTIC SCATTERING

The scattering processes can be subdivided into *elastic* scattering and *inelastic* scattering. The scattering process is said to be elastic if the *total kinetic energy of all the particles before scattering is equal to that after scattering.* In such a process, the internal nature or the structure of the particles does not change and we can apply the law of conservation of energy without taking into consideration the internal energies at all.

But, most of the collisions that occur in nature are inelastic. When two large bodies collide, at least some amount of heat and sound is produced. Thus, some part of the initial kinetic energy is lost in this. This loss must be taken into account while using the law of conservation of energy. When charged particles interact, they may, due to accelerations acquired as a result of the force of interaction, radiate energy in the form of electro-magnetic waves. In some interactions, the masses of the particles may change or new particles may be created or some particles may coalesce into one. Changes are known to take place in the internal structure of the particles, or the total number of particles in the final state may be different from that in the initial state. All such scattering processes or reactions can be grouped, in a broad sense, under inelastic scattering.

7.2 ELASTIC SCATTERING: LABORATORY AND CENTRE OF MASS SYSTEMS

We shall now restrict our attention to elastic collision between two particles. We shall, further, assume that the velocities of the particles are very small as compared to the velocity of light. Hence, the treatment that will follow is non-relativistic. In the general case, both the particles move in such directions that they come closer to each other and collision takes place (Fig. 7.1a). Experiments are often carried out in which one of the particles is at rest in the laboratory and the other approaches it and collision takes place (Fig. 7.1b). Such a set up in which a particle collides with another particle at rest is called the laboratory frame of reference or in short the lab frame.



(a) (b) (c)

Fig. 7.1 Collision between two particles: (a) general; (b) in laboratory frame; and (c) in centre of mass frame

The discussion of collision between two bodies is found to become very simple and symmetric if we describe the effects of the collision in a co-ordinate system in which the centre of mass is at rest. This system is referred to as the centre-of-mass system or in short as the C.M. system. In this system, the collision between two particles is treated as if they have equal and opposite momenta initially (Fig. 7.1c). In that case, we can also say that the 'centre of momentum' of the colliding particles is fixed. If the number of colliding particles is more than two, then in the C.M. system, the vector sum of initial momenta, i.e., $\sum_i \mathbf{k}_i'$ is zero. Then, by the law of conservation of linear momentum, the vector sum of final momenta must also be zero.

The actual measurements in experiments are usually made in the lab system. If we wish to avail of the advantages of the simplifications obtained in a C.M. system, we must have the transformation relations between the quantities measured in the lab system and those measured in the C.M. system and vice versa.

We shall adopt the following notation:

The mass is denoted by m, the initial and final velocities by $\mathbf{u}$ and $\mathbf{v}$, the initial and final momenta by $\mathbf{k}$ and $\mathbf{p}$, the initial and final kinetic energies (non-relativistic) by K and T, respectively, and the angle of scattering by θ. Suffix 1 or 2 will be used to describe the quantities associated with the first or the second particle respectively. In the C.M. frame, the same notation with primes will be used. Further, $M=m_1+m_2$ denotes the total mass and $\mathbf{V}$ denotes the velocity of the centre of mass in the lab system.

Thus, m_1, $\mathbf{u}_1(\mathbf{v}_1)$, $\mathbf{k}_1(\mathbf{p}_1)$ and $K_1(T_1)$ are, respectively, the mass, velocity, momentum and the kinetic energy in the initial (final) state of the first particle.

Similarly, m_2, $\mathbf{u}_2(\mathbf{v}_2)$, $\mathbf{k}_2(\mathbf{p}_2)$ and $K_2(T_2)$ are the corresponding quantities for the second particle.

According to the laws of conservation of linear momentum and kinetic energy, we have

$$\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{p}_1 + \mathbf{p}_2 \tag{7.1}$$

and

$$K_1 + K_2 = T_1 + T_2 \tag{7.2}$$

Equation (7.2) can be written in terms of the corresponding momenta as

$$\frac{k_1^2}{2m_1} + \frac{k_2^2}{2m_2} = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} \tag{7.3}$$

In the collision experiments the initial conditions of the two particles, viz. the masses, the magnitudes of momenta and the trajectories are generally known. This means that we know the six components of $\mathbf{k}_1$ and $\mathbf{k}_2$ together and the ratio of their masses. What is yet to be determined is the six components of final momenta $\mathbf{p}_1$ and $\mathbf{p}_2$ together. But, only four equations—three component equations of (7.1) and (7.2)—are available.

Hence, some additional information, say of direction (spherical angles θ and φ) of one of the scattered particles is necessary. Then, equations (7.1) and (7.2) can be solved to determine the final momentum, say $\mathbf{p}_2$ and the magnitude of $\mathbf{p}_1$. It should be noted that all the final quantities in the final condition (six components of momenta) cannot be determined with the knowledge of initial conditions alone. Some knowledge about motion in the final state is necessary. This is obtained from the experimental observations.

Figure 7.2 describes the elastic collision in the lab and the C.M.

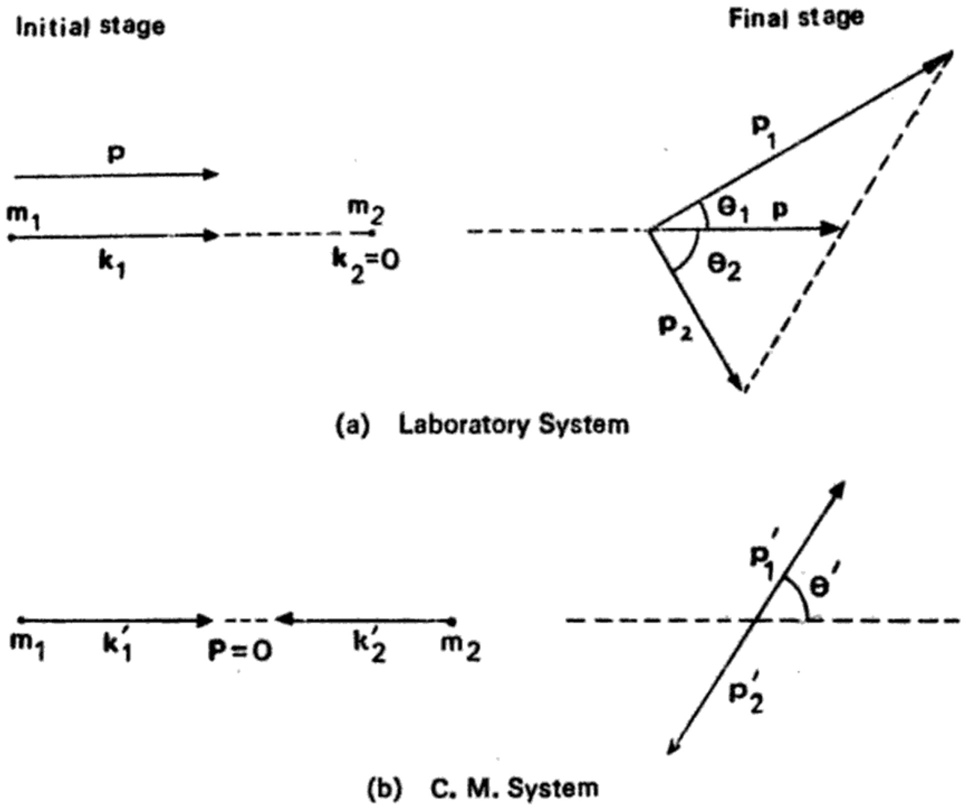


Fig. 7.2 Elastic collision in (a) lab system and (b) centre of mass system. Initial stage in (a) shows particle one with momentum $\mathbf{k}_1$ impinging on particle two at rest with $\mathbf{P}$ as the centre of mass momentum and in (b) two particles approach each other with equal and opposite momenta. Final stages in (a) and (b) show the direction and momenta of the two particles after collision

systems. It should be noted that in the lab system, linear momentum $\mathbf{P}$ of the centre of mass of the two bodies is not zero whereas in the C.M. system it is equal to zero.

The final state of the scattered particles in the lab system has been redrawn in Fig. 7.3. Let $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{R}$ denote the position vectors of the first and the second particles and their centre of mass, respectively, in the lab system.

Let $\mathbf{V}$ be the velocity of the centre of mass so that its linear momentum is

$$\mathbf{P} = M\mathbf{V} = (m_1 + m_2)\mathbf{V}$$

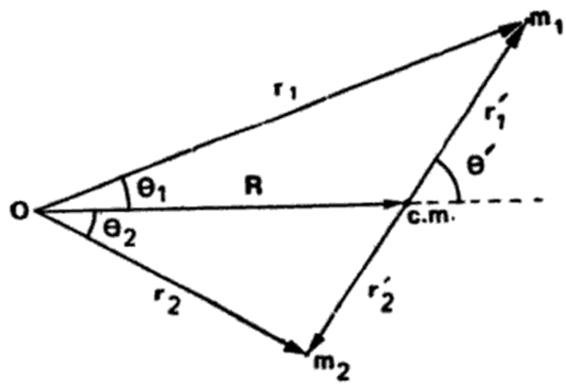


Fig. 7.3 Final state of collision in lab frame showing positions $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{R}$ of particles of mass m_1, m_2 and centre of mass respectively

In the C.M. system, the position vectors of the particles are denoted by $\mathbf{r}_1'$ and $\mathbf{r}_2'$ respectively. Then, the separation between the particles is given by

$$\mathbf{r} = \mathbf{r}_1' - \mathbf{r}_2' = \mathbf{r}_1 - \mathbf{r}_2 \tag{7.4}$$

Since the origin of the C.M. system is situated at the centre of mass, we have

$$m_1\mathbf{r}_1' + m_2\mathbf{r}_2' = 0 \tag{7.5}$$

Adding and subtracting $m_2\mathbf{r}_1'$ from equation (7.5) and using equation (7.4), we get

$$\mathbf{r}_1' = \frac{m_2}{m_1 + m_2}\mathbf{r}$$

or

$$\mathbf{r}_1' = \frac{\mu}{m_1}\mathbf{r} \tag{7.6}$$

where μ is the reduced mass.

Similarly

$$\mathbf{r}_2' = -\frac{\mu}{m_2}\mathbf{r} \tag{7.7}$$

Differentiating equations (7.6) and (7.7) with respect to time, we get the expressions for the velocities in the C.M. system, viz.

$$\dot{\mathbf{r}}_1' = \frac{\mu}{m_1}\dot{\mathbf{r}} \text{ and } \dot{\mathbf{r}}_2' = -\frac{\mu}{m_2}\dot{\mathbf{r}} \tag{7.8}$$

But

$$\dot{\mathbf{r}} = \dot{\mathbf{r}}_1' - \dot{\mathbf{r}}_2' = \dot{\mathbf{r}}_1 - \dot{\mathbf{r}}_2 = \mathbf{u} \tag{7.9}$$

where $\mathbf{u}$ is the relative velocity of the first particle with respect to the second particle. Hence, we get the characteristic property of the C.M. system as

$$m_1\dot{\mathbf{r}}_1' = \mu\mathbf{u} = -m_2\dot{\mathbf{r}}_2' \tag{7.10}$$

Thus, the linear momenta of the particles are equal and opposite both in the initial as well as in the final state.

In the C.M. frame, the laws of conservation of linear momentum and

energy can be written as

$$\mathbf{k}_1' + \mathbf{k}_2' = \mathbf{p}_1' + \mathbf{p}_2' = 0 \tag{7.11}$$

or

$$|\mathbf{k}_1'| = |\mathbf{k}_2'| \quad \text{and} \quad |\mathbf{p}_1'| = |\mathbf{p}_2'|$$

and

$$\tfrac{1}{2}m_1u_1'^2 + \tfrac{1}{2}m_2u_2'^2 = \tfrac{1}{2}m_1v_1'^2 + \tfrac{1}{2}m_2v_2'^2 \tag{7.12}$$

But, from equation (7.10), we have

$$\mathbf{k}_1' = \mu\mathbf{u} = -\mathbf{k}_2'$$

and

$$\mathbf{p}_1' = \mu\mathbf{u} = -\mathbf{p}_2'$$

Hence, equation (7.12) can also be written as

$$\frac{k_1'^2}{2\mu} = \frac{p_1'^2}{2\mu} = \tfrac{1}{2}\mu u^2 \tag{7.13}$$

Thus, we have $u_1' = v_1'$ and $u_2' = v_2'$

This shows that the velocities of the particles in the C.M. system before and after the collision are equal in magnitude.

In the lab system, equations (7.1) and (7.2) assume the form

$$\mathbf{k}_1 = \mathbf{p}_1 + \mathbf{p}_2 \tag{7.14a}$$

and

$$K_1 = T_1 + T_2 \tag{7.14b}$$

or

$$\frac{k_1^2}{2m_1} = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} \tag{7.15}$$

Relations Between Different Quantities in the Laboratory and C.M. Systems

We now obtain the relations between velocities, momenta and the scattering angles in the lab and the C.M. systems. From Fig. 7.3, we can write

$$\mathbf{r}_1 = \mathbf{R} + \mathbf{r}_1' \text{ and } \mathbf{r}_2 = \mathbf{R} + \mathbf{r}_2' \tag{7.16}$$

The corresponding final velocities and momenta are given by

$$\mathbf{v}_1 = \mathbf{V} + \mathbf{v}_1 = \mathbf{V} + \frac{\mu}{m_1}\mathbf{u} \tag{7.17a}$$

and

$$\mathbf{v}_2 = \mathbf{V} + \mathbf{v}_2' = \mathbf{V} - \frac{\mu}{m_2}\mathbf{u} \tag{7.17b}$$

$$\mathbf{p}_1 = m_1\mathbf{v}_1 = m_1\mathbf{V} + m_1\mathbf{v}_1' = m_1\mathbf{V} + \mu\mathbf{u} \tag{7.18a}$$

and

$$\mathbf{p}_2 = m_2\mathbf{v}_2 = m_2\mathbf{V} + m_2\mathbf{v}_2' = m_2\mathbf{V} - \mu\mathbf{u} \tag{7.18b}$$

We can also obtain similar relations between the initial velocities and momenta.

The relation for the momentum of the centre of mass can be obtained from equation (7.18). It is

$$M\mathbf{V} = (m_1 + m_2)\mathbf{V} = \mathbf{p}_1 + \mathbf{p}_2 = \mathbf{k}_1 + \mathbf{k}_2 \tag{7.19}$$

The above relations can be very conveniently represented geometrically and some of the features of the scattering process can be grasped with the aid of the resulting diagrams.

Since the magnitudes of $\mathbf{p}_1'$ and $\mathbf{p}_2'$ are equal, we draw a circle with centre O and radius equal to $|\mathbf{p}_1'| = \mu|\mathbf{u}|$. Then, let us draw vector **AC**

such that vector **AO** represents $m_1\mathbf{V}$ and vector **OC** represents $m_2\mathbf{V}$. Then, by equation (7.18), viz. $\mathbf{p}_1 = m_1\mathbf{V} + \mu\mathbf{u}$ and $\mathbf{p}_2 = m_2\mathbf{V} - \mu\mathbf{u}$ vector **AB** represents $\mathbf{p}_1$ and vector **BC** represents $\mathbf{p}_2$ (Fig. 7.4).

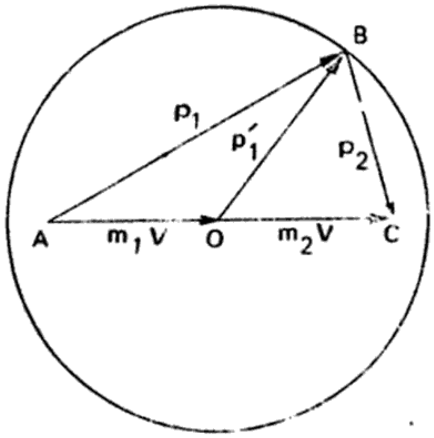


Fig. 7.4 Vector representation of final momenta

In the laboratory system, we know that the initial linear momentum of the particle of mass m_2, viz. $\mathbf{k}_2$ is zero. Hence, equation (7.19) becomes $\mathbf{k}_1 = (m_1 + m_2)\mathbf{V}$. The magnitude of linear momentum represented by vector **OC** is $m_2|\mathbf{V}|$.

But, $$m_2|\mathbf{V}| = m_2 \frac{|\mathbf{k}_1|}{m_1 + m_2}, \text{ from equation (7.19)}$$

$$= \frac{\mu}{m_1}|\mathbf{k}_1| = \mu|\mathbf{u}_1|, \text{ since } \mathbf{k}_1 = m_1\mathbf{u}_1$$

Similarly, vector **OB** represents the linear momentum $\mathbf{p}_1'$ whose magnitude is given by

$$|\mathbf{p}_1'| = \mu|\mathbf{u}|$$

$$= \mu|\mathbf{u}_1 - \mathbf{u}_2|$$

or $$|\mathbf{p}_1'| = \mu|\mathbf{u}_1|, \text{ since } \mathbf{u}_2 = 0 \tag{7.20}$$

Thus, the lengths of vectors **OC** and **OB** are equal. Hence, point C must lie on the circle (Fig. 7.5).

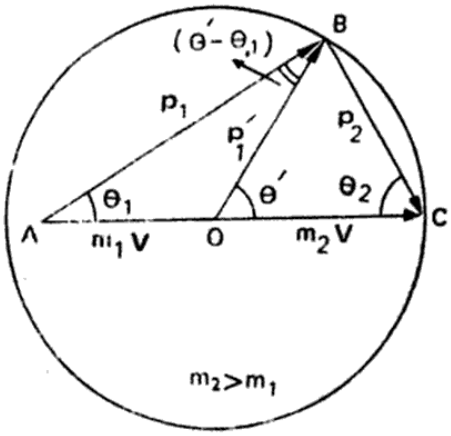


Fig. 7.5 $|m_2\mathbf{V}|$ in the laboratory frame is equal to $|\mathbf{p}_1'|$

The position of point A will be decided by the ratio of the masses of the particles. It is obvious that

$$\frac{AO}{OC}=\frac{m_1}{m_2} \tag{7.21}$$

If $m_2 > m_1$, point A lies inside the circle (Fig. 7.5) while if $m_1 > m_2$, point A lies outside the circle (Fig. 7.6). In case, $m_1 = m_2$, point A will lie on the circle.

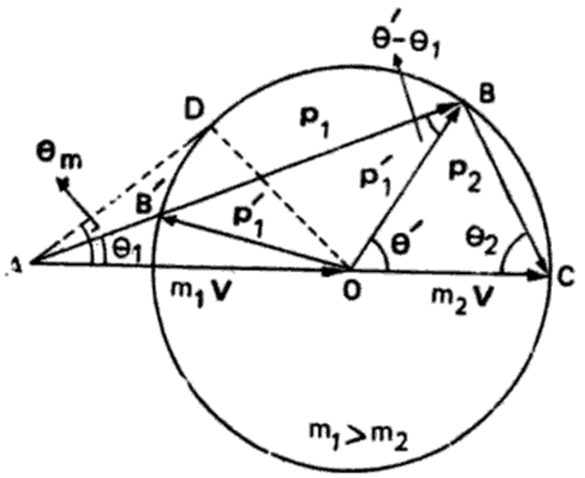


Fig. 7.6 If $m_1 > m_2$, there are two values of final momentum

It should be noted that if $m_2 > m_1$ (Fig. 7.5), for a given ratio of masses and initial momenta, there exists, as seen from the triangle in this figure, only one value of the final momentum represented by vector **AB**. If, however, $m_1 > m_2$ (Fig. 7.6), there exist two values of the final momenta represented by vectors **AB** and **AB**′ for the first particle and by vectors **BC** and **B**′**C** for the second particle. Vectors **AB** and **BC** correspond to forward scattering for which $\theta' < \frac{\pi}{2}$ whereas vectors **AB**′ and **B**′**C** correspond to backward scattering for which $\theta' > \frac{\pi}{2}$. This is as far as the C.M. coordinate system is concerned. In the lab system, scattering angle θ_1 is smaller than $\frac{\pi}{2}$ for both forward and backward scattering. This scattering angle will vary between zero (when $AB = AC$) corresponding to no scattering and θ_{max} when AB is tangential to the circle ($AB = AD$). From Fig. 7.6

$$\sin\theta_{max} = \frac{OD}{OA} = \frac{OC}{OA} = \frac{m_2}{m_1} \tag{7.22}$$

The relations between the scattering angles follow immediately. From $\triangle OBC$, we have

$$2\theta_2 + \theta' = \pi$$

i.e.

$$\theta_2 = \frac{\pi - \theta'}{2} \tag{7.23}$$

where θ_2 is the recoil angle of the second particle.

If $m_1 = m_2$, point A will be on the circle and the angle of scattering will be given by

$$\theta_1 = \tfrac{1}{2}\theta' \quad \text{and} \quad \theta_1 + \theta_2 = \frac{\pi}{2} \tag{7.24}$$

Thus, if $m_1 = m_2$, the particles, after collision, move away at right angles to each other in the lab system. This result is true only in the case of non-relativistic velocities. In this case, θ_1 is the maximum angle of scattering.

From Figs. 7.5 and 7.6, we observe that

$$\tan\theta_1 = \frac{p_1' \sin\theta'}{m_1 V + p_1' \cos\theta'} = \frac{\sin\theta'}{\cos\theta' + \dfrac{m_1 V}{p_1'}} \tag{7.25}$$

But

$$\frac{m_1 V}{p_1'} = \frac{m_1}{m_2} \tag{7.26}$$

Hence

$$\tan\theta_1 = \frac{\sin\theta'}{\cos\theta' + \dfrac{m_1}{m_2}} \tag{7.27a}$$

If now, $m_1 = m_2$

$$\tan\theta_1 = \frac{\sin\theta'}{\cos\theta' + 1} = \tan\frac{\theta'}{2}$$

or $\theta_1 = \dfrac{\theta'}{2}$, which is the same as equation (7.24). From Figs. 7.5 and 7.6, we can relate recoil angle θ_2 with the C.M. angle θ'. Thus

$$\tan\theta_2 = \frac{p_1' \sin\theta'}{m_2 V - p_1' \cos\theta'} = \frac{\sin\theta'}{\dfrac{m_2 V}{p_1'} - \cos\theta'}$$

$$= \frac{\sin\theta'}{1 - \cos\theta'} = \cot\frac{\theta'}{2}$$

since $\dfrac{m_2 V}{p_1'} = 1$. Thus we get

$$\theta_2 = \frac{\pi - \theta'}{2}$$

the relation already obtained.

We can solve equation (7.27a) to obtain θ' in terms of θ. Thus, squaring relation (7.27a) and expressing $\sin\theta'$ in terms of $\cos\theta'$, we get a quadratic equation in

$$\cos^2\theta' + 2\left(\frac{m_1}{m_2}\right)\sin^2\theta_1 \cos\theta' + \left(\frac{m_1}{m_2}\right)^2 \sin^2\theta_1 - \cos^2\theta_1 = 0$$

Hence, corresponding to every value of θ_1, we have two values of θ' given by

$$\cos\theta'_{1,2} = -\left(\frac{m_1}{m_2}\right)\sin^2\theta_1 \pm \cos\theta_1\sqrt{1 - \left(\frac{m_1}{m_2}\right)^2 \sin^2\theta_1} \tag{7.27b}$$

For $m_2 > m_1$, it is clear from Fig. 7.5 that there is only one value of θ' for every θ_1 and $0 < \theta_1 < \pi$ and we have to choose the value with positive sign. (Because when $\theta_1 = 0$ we should get $\theta' = 0$, a case of no scattering.)

For $m_1 > m_2$, Fig. 7.6 clearly indicates that there are two values of θ' for every θ_1, and $0 < \theta_1 < \theta_{max}$. As θ_1 increases, one of the values of θ' increases while the other decreases.

From equation (7.27b), when $\sin \theta_{1_{max}} = \dfrac{m_2}{m_1}$, we get $\cos \theta' = -\sin \theta_{1_{max}}$ or the maximum value of the angle of scattering is

$$\theta_{1_{max}} = \theta' - \frac{\pi}{2}$$

For the scattering of particles of equal masses, as θ' varies between 0 and π, θ_1 varies between 0 and $\frac{\pi}{2}$ and θ_2 varies between $\frac{\pi}{2}$ and 0, keeping $\theta_1 + \theta_2 = \frac{\pi}{2}$ constant.

7.3 KINEMATICS OF ELASTIC SCATTERING IN THE LABORATORY SYSTEM

Consider the elastic scattering of two particles in the laboratory system (Fig. 7.7). We rewrite equations (7.14) and (7.15) which express the laws

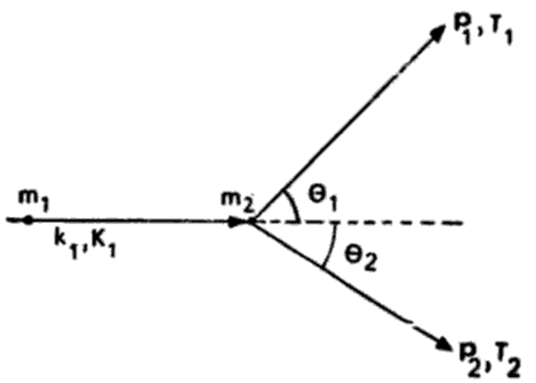


Fig. 7.7 Elastic scattering of two particles in the laboratory frame

of conservation of linear momentum and energy. These are

$$\mathbf{k}_1 = \mathbf{p}_1 + \mathbf{p}_2 \tag{7.14}$$

and

$$K_1 = T_1 + T_2$$

or

$$\frac{k_1^2}{2m_1} = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} \tag{7.15}$$

Let the particle of mass m_1 be scattered through an angle θ_1. We shall now obtain the expressions for final momenta and energies in terms of θ_1. For this, we eliminate p_2 from equations (7.14) and (7.15) as follows:

Squaring equation (7.14), we get

$$p_2^2 = k_1^2 + p_1^2 - 2k_1 p_1 \cos \theta_1 \tag{7.28}$$

Now, from equation (7.15)

$$\frac{p_2^2}{2m_2} = \frac{k_1^2}{2m_1} - \frac{p_1^2}{2m_1}$$

i.e.
$$\frac{k_1^2 + p_1^2 - 2k_1 p_1 \cos\theta_1}{2m_2} = \frac{k_1^2}{2m_1} - \frac{p_1^2}{2m_1}$$

or
$$p_1^2\left(\frac{1}{m_1} + \frac{1}{m_2}\right) - p_1\frac{2k_1\cos\theta_1}{m_2} - k_1^2\left(\frac{1}{m_1} - \frac{1}{m_2}\right) = 0$$

or
$$p_1^2\left(\frac{m_1 + m_2}{m_1 m_2}\right) - p_1\frac{2k_1\cos\theta_1}{m_2} + k_1^2\left(\frac{m_1 - m_2}{m_1 m_2}\right) = 0 \tag{7.29}$$

This is a quadratic in $\frac{p_1}{k_1}$ and has a solution

$$\frac{p_1}{k_1} = \frac{m_1}{m_1 + m_2}\cos\theta_1 \pm \sqrt{\left(\frac{m_1}{m_1 + m_2}\right)^2\cos^2\theta_1 - \left(\frac{m_1 - m_2}{m_1 + m_2}\right)} \tag{7.30}$$

Here, k_1 and θ_1 are assumed to be known. Then, using equation (7.30), we can evaluate p_1. Once p_1 is known, we can use equation (7.28) to evaluate p_2. Further, the angle of recoil of the second particle, viz. θ_2 can be found out by solving equation (7.14).

If $m_2 > m_1$, the quantity under the radical sign in equation (7.30) is always positive. If $m_1 > m_2$, the quantity under the radical sign will become zero, if $\theta = \theta_{max}$, and is given by

$$\cos^2\theta_{max} = 1 - \frac{m_2^2}{m_1^2}, \quad 0 \leqslant \theta_{max} \leqslant \frac{\pi}{2} \tag{7.31}$$

or
$$\sin\theta_{max} = \frac{m_2}{m_1} \tag{7.32}$$

Thus, for $\theta_1 > \theta_{max}\,(\theta_1 < \pi)$, the quantity $\frac{p_1}{k_1}$ is imaginary and hence θ_{max} is the maximum angle of scattering allowed physically. If the mass of the incident particle is large, i.e., if $m_1 \gg m_2$, $\cos\theta_{max} = 1$ or $\theta_{max} = 0$. Thus, such a heavy incident particle hardly deviates from its original path. These results can be very well understood with the help of Fig. 7.6. Thus, for $\theta_1 < \theta_{max}$, and $m_1 > m_2$, we have two values of p_1 as obtained from equation (7.30). These were predicted earlier by Fig. 7.6.

In the above treatment, the final direction of the particle is determined only if angle θ (the polar angle—if the direction of $\mathbf{k}_1$ is taken along the z-axis) is known. The azimuthal angle (φ) does not appear. This means that we have assumed azimuthal symmetry in the scattering process.

Loss of Kinetic Energy

Since the velocities of a particle before and after the collision in the C.M. system are the same and the mass of the particle is constant, there is no change in the kinetic energy of the particle before and after collision.

In the laboratory system, the ratio of kinetic energy after and before

the collision is given by

$$\frac{T_1}{K_1} = \frac{\frac{1}{2}m_1v_1^2}{\frac{1}{2}m_1u_1^2} = \frac{p_1^2}{k_1^2}$$

Substituting the value of $\frac{p_1}{k_1}$ from equation (7.30) and simplifying it, we get

$$\frac{T_1}{K_1} = \frac{m_1^2}{(m_1 + m_2)^2}\left[\cos\theta_1 \pm \sqrt{\frac{m_2^2}{m_1^2} - \sin^2\theta_1}\right]^2 \tag{7.33}$$

Out of the two signs, the positive sign is chosen before the quantity under the radical sign. This follows from the fact that on scattering $T_1 = K_1$ and $\theta_1 = 0$ and the two sides of equation (7.33) should match. If $m_1 > m_2$, the expression (7.33) is double-valued and ratio $\frac{T_1}{K_1}$ has a maximum value corresponding to the angle of scattering θ_{max}, in which case the term under the radical sign vanishes.

Now, by the law of conservation of energy, the ratio of the kinetic energy of the recoiled particle to that of the incident particle is

$$\frac{T_2}{K_1} = 1 - \frac{T_1}{K_1} \tag{7.34}$$

If $m_1 = m_2$, formulae (7.33) and (7.34) assume simple forms, viz.

$$\frac{T_1}{K_1} = \cos^2\theta_1 \quad \text{and} \quad \frac{T_2}{K_1} = \sin^2\theta_1 \tag{7.35}$$

Thus, if $\theta_1 = \frac{\pi}{2}$, or $\theta' = \pi$, i.e., in the case of backward scattering, $\frac{T_1}{K_1} = 0$. The incident particle, therefore, loses all its kinetic energy. It is completely imparted to the recoiled particle.

This last result finds application in nuclear reactors. It is well known that in a reactor, certain substances are used as moderators. Their function is to slow down the thermal neutrons liberated in the nuclear reaction. An ideal moderator is one which has an atomic mass equal to the mass of the neutron. However, because of the other considerations of nuclear reactions, either heavy water (which contains deuterium with mass number 2) or graphite are used for this purpose. In the laboratory experiments, hydrogen in the form of paraffin hydrocarbons is mostly used for slowing down the neutrons.

7.4 INELASTIC SCATTERING

Particles such as nuclei, atoms and molecules possess some kind of structure, i.e., they are formed of a number of constituent particles arranged in some order. In such cases, electromagnetic or nuclear fields are present and the constituent particles move in these fields in bounded motion. Such particles, therefore, possess internal energies both potential and kinetic. Whenever these particles participate in scattering experiments, the internal energy may also change. Such a scattering process has been termed as inelastic scattering. When a beam of protons is

incident on nuclei, inelastic scattering can take place in many ways. A part of the kinetic energy of the incident proton may be absorbed by the nucleus, the proton may be absorbed and a neutron along with some new particle may be released, etc.

Inelastic collisions are divided into two types: (i) *endoergic*, in which case translational kinetic energy is absorbed, and (ii) *exoergic*, in which case kinetic energy is released in the process of scattering. It should be noted that we are not considering rotational kinetic energy of the particles.

In the case of inelastic collisions, the law of conservation of linear momentum holds good. Consider a case of collision, in which a particle of mass m_1 collides with a particle of mass m_2 initially at rest. After collision, these particles fly apart as different particles of masses m_3 and m_4, respectively. The process as in the lab system is shown in Fig. 7.8.

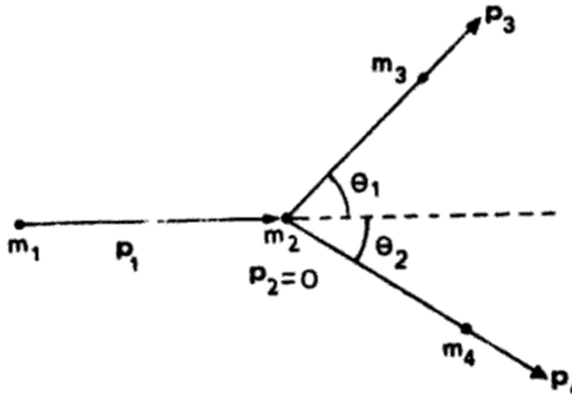


Fig. 7.8 Inelastic collision in the laboratory frame

Applying the laws of conservation of linear momentum and energy, we get

$$\mathbf{p}_1 = \mathbf{p}_3 + \mathbf{p}_4 \tag{7.36}$$

$$T_1 = T_3 + T_4 + Q \tag{7.37}$$

or

$$\frac{p_1^2}{2m_1} = \frac{p_3^2}{2m_3} + \frac{p_4^2}{2m_4} + Q \tag{7.38}$$

Quantity Q in equations (7.37) and (7.38) is the amount of kinetic energy absorbed or released in the scattering process.

(i) If $Q < 0$, the final kinetic energy of the two particles is greater than their initial kinetic energy. Hence, $T_3 + T_4 > T_1$. Thus, in this process the energy is released and the process is *exoergic*.

(ii) If $Q > 0$, the final kinetic energy of the two particles is less than their initial kinetic energy. Hence, $T_3 + T_4 < T_1$. Thus, in this process the energy is absorbed and the process is *endoergic*.

(iii) If $Q = 0$, the process of scattering is an *elastic* one.

In many scattering experiments, momentum $\mathbf{p}_1$ of the first particle is known after the collision. The scattering angle of this particle, viz. θ_1, is recorded in the experiment. We shall, therefore, eliminate p_4 from

the above equations and express the final result in terms of θ_1. On squaring equation (7.36), we get

$$p_4^2 = p_1^2 + p_3^2 - 2p_1p_3 \cos \theta_1 \tag{7.39}$$

Substituting this value in equation (7.38), we get

$$\begin{aligned} Q &= \frac{p_1^2}{2m_1} - \frac{p_3^2}{2m_3} - \frac{p_1^2 + p_3^2 - 2p_1p_3 \cos \theta_1}{2m_4} \\ &= \frac{p_1^2}{2m_1}\left(1 - \frac{m_1}{m_4}\right) - \frac{p_3^2}{2m_3}\left(1 + \frac{m_3}{m_4}\right) + \frac{2p_1p_3 \cos \theta_1}{2m_4} \\ &= T_1\left(1 - \frac{m_1}{m_4}\right) - T_3\left(1 + \frac{m_3}{m_4}\right) + 2\sqrt{\frac{T_1T_3m_1m_3}{m_4^2}} \cos \theta_1 \end{aligned} \tag{7.40}$$

Equation (7.40) is very often used in nuclear physics for finding out the Q-values in nuclear reactions. The Q-value decides whether a particular scattering process—called reaction—is energetically possible or not. The condition for this is

$$Q + T_1 \geqslant 0$$

This is satisfied by exoergic reaction. In the case of endoergic reactions, however, the incident particle must have some minimum energy called the *threshold energy* to produce particles having masses m_3 and m_4.

In the laboratory system, the law of conservation of linear momentum requires that the final particles should carry the momentum of the incident particle and hence must have the corresponding energy. In other words, the kinetic energy of the centre of mass is not available for the reaction to take place. At the threshold, for endoergic reactions, the final particles move out with the energy of the centre of mass.

In the C.M. system, the centre of mass is at rest and hence all the kinetic energy of the incident particles is available for reaction.

7.5 CROSS-SECTION

We have so far obtained the relation between the velocities or momenta of the colliding particles in the initial state to those in the final state. Such collisions of two large bodies or the 'classical' particles can be observed and the results derived can be applied to these. However, we could never predict the value of the scattering angle since the nature of the law of force was not known. We have already mentioned that the scattering experiments form an important method by which we can investigate the nature of the force-field existing between the microparticles.

In collision experiments, particularly of microparticles (atomic and subatomic particles), we cannot follow the individual particle through the scattering process. Also, particles involved in scattering experiments are identical and cannot be distinguished. This type of situation cannot be tackled through the basic concepts of classical mechanics which is deterministic. New concept of scattering probability is, therefore, introduced through 'cross-section'. It is very useful in the study of the inter-

actions of microparticles in modern physics. Moreover, the kinematical formulae developed for collisions of particles are directly useful even in the case of scattering of microparticles.

The problem of scattering of microparticles will be dealt in an essentially different way. A beam of such microparticles is made incident on the target particles under investigation. The scattered particles at some scattering angle are recorded. We can measure the velocities or energies of the scattered particles. Since we cannot 'see' the phenomenon that takes place at the time of interaction, we call the region of interaction the 'black-box'. But, from the knowledge of the distribution of the scattered particles, we can try to know the nature of the force that produces the observed distribution of the scattered particles.

Since we are dealing not with a single particle but with beams of incident and scattered particles, we have to speak in terms of the probability of scattering of some of the particles in some specified direction. The probability is expressed in terms of the *differential cross-section.*

While dealing with scattering problems in classical as well as quantum mechanics, we consider the interaction between a single incident particle and a single target particle. But, the results obtained therefrom are used for a beam of particles incident on a large collection of target particles. In doing so, we make an implicit assumption that the incident as well as the target particles are sufficiently apart so that the process of scattering of one incident particle by one target particle is unaffected by the presence of other particles either in the beam or in the target. Further, the same type of interaction is present in all pairs of incident and target particles. These assumptions are found to be valid in most of the experiments. For example, in the Rutherford scattering experiment, the target is in the form of a thin foil and any alpha particle from the incident beam experiences the same type of Coulomb field when it approaches the nucleus. Moreover, we assume that an alpha particle normally suffers a single scattering as a result of its interaction with one of the nuclei of the foil. The probability of scattering of an alpha particle by two nuclei is very much smaller and hence it is negligible.

In the case of scattering of the two particles, the force of mutual interaction is responsible for producing the deflection of particles from their original path. We can treat this problem as an equivalent one-body problem, i.e., a single particle of mass μ being scattered by a fixed central force-field of interaction. Thus considered, the scattering is in the C.M. system. The results obtained in Chapter 5 for unbounded motion can be applied to this problem.

We shall now define the differential cross-section in the C.M. system.

Let N be the number of particles incident per unit area per second on target particles which are at rest. Quantity N is very often called the intensity of the incident beam. Let dN be the number of particles scattered per second in solid angle $d\Omega'$ along a direction given by θ' and φ.

The angle is measured from the direction of incident particles which is taken as the z-axis. The number of particles scattered (dN) in solid angle $d\Omega'$ will obviously be proportional to the intensity of the incident particles and also the magnitude of solid angle. Thus

$$dN \propto N\, d\Omega'$$

or
$$dN = \sigma(\Omega') N\, d\Omega' \tag{7.41}$$

where $\sigma(\Omega')$ is the proportionality constant which will depend on the direction, i.e., on θ' and φ'. The constant is known as the *differential cross-section* and can be written as

$$\sigma(\Omega') \equiv \frac{d\sigma}{d\Omega'} = \frac{1}{N}\frac{dN}{d\Omega'} \tag{7.42}$$

It should be noted that solid angle $d\Omega'$ is given by

$$d\Omega' = \sin\theta'\, d\theta'\, d\varphi' \tag{7.43}$$

Since, the azimuthal symmetry is assumed in the scattering process, the dependence on φ' is absent and the integration over φ' gives factor 2π. Hence we can choose solid angle as

$$d\Omega' = 2\pi \sin\theta'\, d\theta' \tag{7.44}$$

The name cross-section is justified since the dimension of $\sigma(\Omega')$ is that of area. The differential cross-section is a measure of the probability that a particle will be scattered in solid angle $d\Omega'$ along direction (θ', φ'). This can also be conceived as the *effective area* posed by the scatterer to the incident particle.

The scattering in the laboratory and C.M. systems is shown in Fig. 7.9.

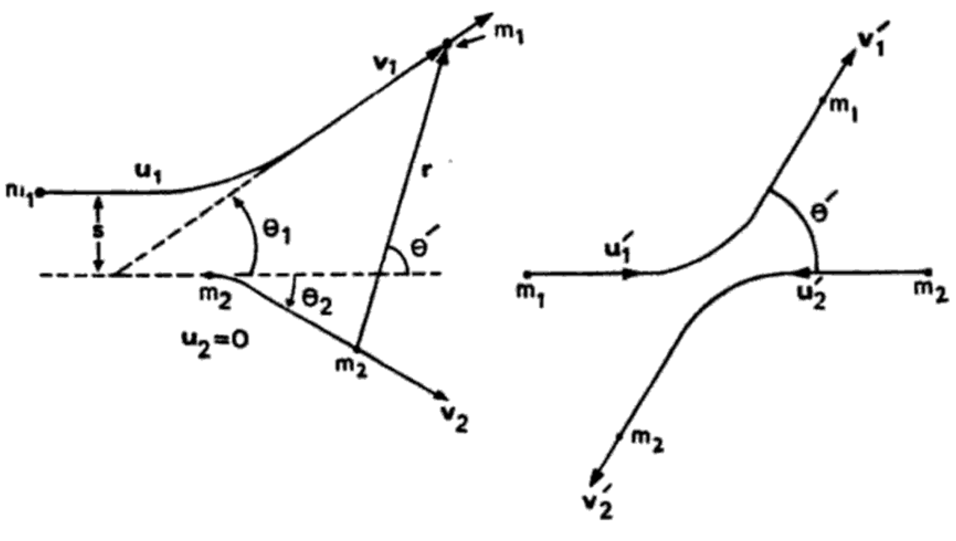


Fig. 7.9 Scattering in the: (a) laboratory system; (b) centre-of-mass system

Let $\mathbf{r}$ be the relative position vector (i.e., a vector joining the position of particle 2 to that of 1) of the two particles. As seen from (Fig. 7.9b), it will be parallel to $\mathbf{u}_2'$ before scattering and parallel to $\mathbf{v}_1'$ after scattering. In the equivalent one-body problem as is proved in article 5.1, $\mathbf{r}$

represents the position of the equivalent single body of mass μ from the centre of force. The corresponding scattering is shown in Fig. 7.10.

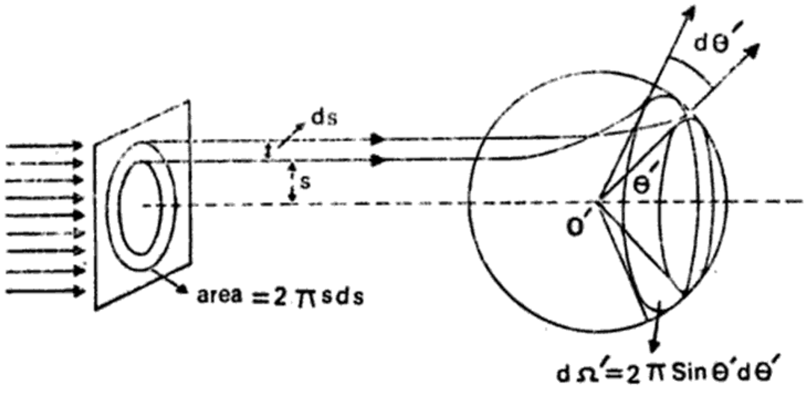


Fig. 7.10 Scattering of particles by a centre of force O'

Consider a particle in a beam of particles incident on a scattering centre which is treated as fixed. If the particle is not deflected, it will miss the target by distance s. This distance is called the *impact parameter*. This particle is scattered at angle θ' and has a symmetry about the incident direction, i.e., the azimuthal symmetry. A particle with impact parameter $s + ds$ will obviously be scattered through a smaller angle $(\theta' - d\theta')$. All the particles in the annular ring of area $2\pi s\, ds$ will be scattered in solid angle $d\Omega' = 2\pi \sin\theta'\, d\theta'$. Hence, the number of particles scattered in this solid angle is

$$dN = \sigma(\theta')N\, d\Omega' = \sigma(\theta')N 2\pi \sin\theta'\, d\theta' \tag{7.45}$$

This number must also be equal to $2\pi s\, ds\, N$.

Hence, we get

$$2\pi s\, ds\, N = -2\pi \sin\theta'\, d\theta' N\sigma(\theta') \tag{7.46}$$

The negative sign is attached to indicate that as impact parameter s increases, scattering angle θ' decreases. Simplifying equation (7.46), we get

$$\sigma(\theta') \equiv \frac{d\sigma(\theta')}{d\Omega'} = \frac{s}{\sin\theta'}\left|\frac{ds}{d\theta'}\right| \tag{7.47}$$

This relation between s and θ' will help us to calculate the differential cross-section. Since cross-section indicates probability, the negative sign is meaningless and hence it is dropped out.

The total cross-section can be obtained by integrating equation (7.47) over all the angles and is given by

$$\sigma_T = \int \sigma(\Omega')\, d\Omega'$$

or

$$\sigma_T = 2\pi \int \sigma(\theta') \sin\theta'\, d\theta' \tag{7.48}$$

The integration on the right-hand side of equation (7.48) extends over all possible solid angles. *The total cross-section represents the number of*

particles scattered in all directions per unit intensity of incident beam per second.

Relation Between Cross-Sections in the C.M. and Laboratory Systems

So far, the problem of scattering of particles was considered in the C.M. system. We will now consider the transformation of the cross-section from the C.M. system to the laboratory system.

Since the number of particles scattered in a solid angle is $d\Omega = 2\pi \sin\theta_1\, d\theta_1$ in the laboratory system it must be equal to that scattered in the corresponding solid angle $d\Omega' = 2\pi \sin\theta'\, d\theta'$ in the C.M. system, we have

$$N\sigma(\theta')2\pi \sin\theta'\, d\theta' = N\sigma(\theta_1)2\pi \sin\theta_1\, d\theta_1 \tag{7.49}$$

or

$$\sigma(\theta') = \sigma(\theta_1)\frac{\sin\theta_1\, d\theta_1}{\sin\theta'\, d\theta'}$$

$$= \sigma(\theta_1)\frac{d(\cos\theta_1)}{d(\cos\theta')} \tag{7.50}$$

The relation between θ_1 and θ' is expressed in equation (7.25). From triangle ABC in Figs. 7.5 and 7.6, using the sine law, we have

$$\frac{\sin(\theta' - \theta_1)}{\sin\theta_1} = \frac{m_1 V}{p_1'} = \frac{m_1}{m_2} \tag{7.51}$$

Differentiation of equation (7.51) gives

$$\frac{d\theta'}{d\theta_1} = 1 + \frac{\sin(\theta' - \theta_1)\cos\theta_1}{\cos(\theta' - \theta_1)\sin\theta_1} \tag{7.52}$$

$$= \frac{\sin\theta'}{\cos(\theta' - \theta_1)\sin\theta_1} \tag{7.53}$$

Hence

$$\sigma(\theta') = \sigma(\theta_1)\frac{\sin^2\theta_1\cos(\theta' - \theta_1)}{\sin^2\theta'} \tag{7.54}$$

The right-hand side of equation (7.54) can be expressed completely in terms of θ_1 by virtue of equation (7.51).

When the masses of the two colliding particles are equal, i.e., $m_1 = m_2$, then $\theta_1 = \frac{\theta'}{2}$ and relation (7.54) becomes

$$\sigma(\theta') = \frac{\sigma(\theta_1)}{4\cos\theta_1} \tag{7.55}$$

When $m_2 \gg m_1$, the laboratory system reduces to a system of scattering of m_1 by m_2 which remains fixed because of its very large mass, i.e., the lab system reduces to a C.M. system and $\theta_1 \approx \theta'$. Hence the cross-section in the laboratory and the centre-of-mass systems are equal. Thus, we have, when $m_2 \gg m_1$

$$\sigma(\theta') = \sigma(\theta_1) \tag{7.56}$$

Consider an example of elastic scattering of spheres each of mass m_1 and radius b from target spheres each of mass m_2 and radius a in the

C.M. and laboratory systems. We shall first obtain differential and total cross-sections in the C.M. system and then relate them to those in the laboratory system.

The law of force in this case is $V = \infty$ for $r < a + b$ and $V = 0$ for $r > a + b$, where V represents the potential. Thus, the incident spheres will get scattered after rebounding from the surface of the target spheres as shown in Fig. 7.11. Before and after the collision, the spheres will move freely. The impact parameter in this case is

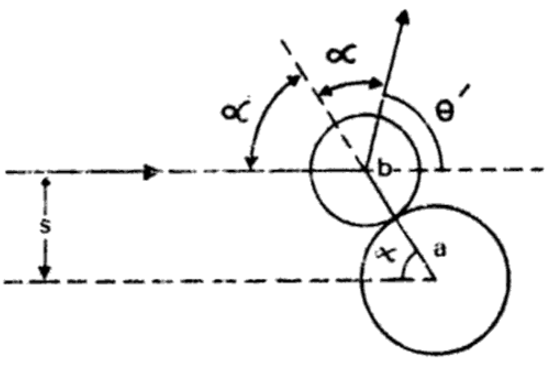


Fig. 7.11 Elastic scattering of two spheres

$$s = (a + b) \sin \alpha = (a + b) \sin \frac{\pi - \theta'}{2} = (a + b) \cos \frac{\theta'}{2}$$

Therefore

$$\frac{ds}{d\theta'} = \frac{a + b}{2} \sin \frac{\theta'}{2} \tag{7.57}$$

From formula (7.47), we get the differential cross-section in the C.M. system as

$$\sigma(\theta') \equiv \frac{d\sigma(\theta')}{d\Omega'} = \frac{(a + b)^2}{4} \tag{7.58}$$

Thus, the scattering of spheres in the C.M. system is isotropic, i.e., it does not depend on θ'. Here, $d\Omega' = \sin \theta' \, d\theta' \, d\varphi'$. After integration we get the total cross-section in the C.M. system

$$\sigma_{CM} = \int \sigma(\theta') \, d\Omega' = \pi(a + b)^2$$

This is the effective area of cross-section and the centres of the two colliding spheres will never be within the distance of $a + b$. If spheres of radius b are point particles then the total cross-section πa^2 is the cross-sectional area of target sphere.

In order to obtain differential cross-section of scattering of spheres of mass m_1 in the laboratory system, we use equation (7.50). Differentiating equation (7.27a), we get

$$-d(\cos \theta') = \sin \theta' \, d\theta' = \sin \theta_1 \, d\theta_1 \left\{ 2 \frac{m_1}{m_2} \cos \theta_1 \pm \frac{1 + \left(\frac{m_1}{m_2}\right)^2 \cos 2\theta_1}{\sqrt{1 - \left(\frac{m_1}{m_2}\right)^2 \sin^2 \theta_1}} \right\}$$

Thus, for $m_2 > m_1$, the differential cross-section in the laboratory system is from equations (7.50) and (7.58)

$$\sigma(\theta_1) = \frac{(a+b)^2}{4}\left\{2\frac{m_1}{m_2}\cos\theta_1 + \frac{1+\left(\frac{m_1}{m_2}\right)^2\cos 2\theta_1}{\sqrt{1-\left(\frac{m_1}{m_2}\right)^2\sin^2\theta_1}}\right\} \tag{7.59}$$

where we have taken only one value of θ' corresponding to positive sign. For $m_1 > m_2$, however, we have to account for both the values of θ' for a given θ_1. As θ_1 increases, one of the corresponding values of θ' increases whereas the other decreases. Hence it is necessary to take the difference and not the sum of the two values of $d(\cos\theta')$ with opposite signs before the radical sign. Thus

$$d(\cos\theta_1') - d(\cos\theta_2') = 2\,d(\sin\theta_1)\frac{1+\left(\frac{m_1}{m_2}\right)^2\cos 2\theta_1}{\sqrt{1-\left(\frac{m_1}{m_2}\right)^2\sin^2\theta_1}}$$

Corresponding to this the cross-section in the lab system is

$$\sigma(\theta_1) = \frac{(a+b)^2}{2}\frac{1+\left(\frac{m_1}{m_2}\right)^2\cos 2\theta_1}{\sqrt{1-\left(\frac{m_1}{m_2}\right)^2\sin^2\theta_1}} \tag{7.60}$$

For $m_1 = m_2$, however, both expressions (7.59) and (7.60) reduce to

$$\sigma(\theta_1) = (a+b)^2|\cos\theta_1| \tag{7.61}$$

Here only positive values of the cross-section are to be taken.

The total cross-section is obtained by integrating all the positive values of equation (7.61) and is

$$\begin{aligned}\sigma_{\text{lab}} &= \int \sigma(\theta_1)\sin\theta_1\,d\theta_1\,d\varphi_1\\ &= 2\pi(a+b)^2\int|\cos\theta_1|\,d(\cos\theta_1)\\ &= \pi(a+b)^2\end{aligned}$$

where the integration is carried between 0 and $\frac{\pi}{2}$, the allowed values of θ_1.

Differential cross-section for recoil spheres can be written from equation (7.58), by using $\theta' = \pi - 2\theta_2$, as

$$\sigma(\theta_2) = \frac{d\sigma(\theta_2)}{d\Omega_2} = (a+b)^2|\cos\theta_2|$$

where $d\Omega_2 = \sin\theta_2\,d\theta_2\,d\varphi_2$

7.6 THE RUTHERFORD FORMULA

In order to calculate the differential cross-section for scattering of two charged particles, we shall first obtain a relation between impact parameter s and scattering angle θ'. The trajectory of the scattered particle—

Substituting the value of ϵ, we get

$$\tan\frac{\theta'}{2} = \left(\frac{\mu\mathcal{K}^2}{2EL^2}\right)^{1/2} = \frac{\mathcal{K}}{s\mu u_1'^2} \tag{7.65}$$

Differentiating equation (7.65), we get the relation between θ' and s namely

$$\frac{d\theta'}{2\cos^2\frac{\theta'}{2}} = \frac{-\mathcal{K}\,ds}{s^2\mu u_1'^2}$$

or

$$\left|\frac{ds}{d\theta'}\right| = \frac{s^2\mu u_1'^2}{2\mathcal{K}\cos^2\frac{\theta'}{2}} \tag{7.66}$$

Substituting the value of s as given by equation (7.65), and that of $\left|\frac{ds}{d\theta'}\right|$ as given by equation (7.66), the differential cross-section of (7.47) becomes

$$\sigma(\theta') = \frac{s}{\sin\theta'}\left|\frac{ds}{d\theta'}\right|$$

$$= \left(\frac{\mathcal{K}}{2\mu u_1'^2}\right)^2 \frac{1}{\sin^4\frac{\theta'}{2}} = \frac{\mathcal{K}^2}{4E^2}\frac{1}{\sin^4\frac{\theta'}{2}} \tag{7.67}$$

The dependence of $\sigma(\theta')$ on $\mathcal{K}^2$ shows that the formula is applicable to the field of attractive force also.

In the case of scattering experiments of Rutherford

$$|\mathcal{K}| = Z_1Z_2e^2 \tag{7.68}$$

where Z_1e and Z_2e are the charges of the incident particle and the target nuclei respectively. It should be noted that e is the magnitude of the electronic charge.

Thus, the differential cross-section is given by

$$\sigma(\theta') = \left(\frac{Z_1Z_2e^2}{2E}\right)^2 \operatorname{cosec}^4\frac{\theta'}{2} \tag{7.69}$$

Equation (7.69) gives the differential cross-section for the Rutherford experiment of scattering of alpha particles ($Z_1 = 2$) from the atomic nuclei.

The *distance of the closest approach* (r_1) is the distance of the point at which the scattered particle turns away from the scattering centre. From Fig. 7.12

$$r_1 = a(1+\epsilon)$$

Taking $a = \frac{\mathcal{K}}{2E} = \frac{Z_1Z_2e^2}{2E}$ and substituting we get

$$r_1 = \frac{Z_1Z_2e^2}{2E}\left[1 + \sqrt{1 + \frac{2EL^2}{Z_1^2Z_2^2e^4\mu}}\right] \tag{7.70}$$

This distance is smallest when impact parameter s and hence angular momentum L are zero. This corresponds to the case of a head-on

collision in the lab system. The minimum value of the distance of closest approach is given by

$$r_{1_{\min}} = r_0 = \frac{Z_1 Z_2 e^2}{E}$$

or

$$E = \frac{Z_1 Z_2 e^2}{r_0} \tag{7.71}$$

Rutherford was able to show that distance r_0 can never be less than 10^{-14} m from his observations in the scattering experiments. From these observations, he concluded that the positive charge of the atom was concentrated within a small region having dimensions of the order of 10^{-14} m. This he called the nucleus.

It should be noted that the C.M. system is identical to the laboratory system when the target particle is very heavy and the recoil is negligible. If the nuclei of the thin metal foil used as a target in the scattering experiment are heavy, the above results are also the results for the particle scattering in the laboratory system.

If we integrate the differential cross-section expressed in equation (7.69), we get a divergent result. The physical reason for such a result is that the Coulomb field of the charged target particle has, in principle, an infinite range. Thus, the incident particles even with very large impact parameters will always be deflected by some angle; maybe a very small angle. In the process of integration carried out for finding out the total cross-section, we add all such contributions. Hence, the small but definite contributions even for larger impact parameters lead us to an indefinite value for the total cross-section. However, all fields having an infinite range do not yield divergent total cross-section. If the potential falls off faster than $\frac{1}{r^2}$ at greater distances, then the total cross-section comes out to be finite.

In the case of atoms, the nuclear Coulomb field is screened by the electrons around it and the field has a finite range. Such a field is represented by screened potential

$$V(r) = \frac{1}{r} e^{-r/a} \tag{7.72}$$

where a is called the screening radius.

A potential of this type yields finite value for the total cross-section. The results obtained above are similar to the results obtained on the basis of quantum theory. Hence, Rutherford's conclusions were applicable even to the micro-particles for which the classical theories prove to be inadequate.

QUESTIONS

1. Explain how the conservation of momentum applies to a handball bouncing off a wall.
2. When dealing with atoms, nuclei or elementary particles, what does

it mean to say that two such bodies 'touch' during a collision?

3. Could we determine, in principle, the cross-section by using only one projectile particle and one target particle? In practice?
4. Explain how Rutherford was led to conclude that the positive charge of the atom is concentrated within a small region of dimensions 10^{-14} m.
5. In case of nuclear reactors, wherein moderators are used to slow down thermal neutrons, which condition is satisfied regarding the masses of the atoms of the moderator and the neutrons?
6. While considering the motion of moon around the earth, can we treat the earth to be stationary? Explain.
7. Will the Rutherford experiment of scattering of α-particle give correct size of the nucleus? Why?
8. Assuming that the densities of the moon and the earth are equal, to what extent are we justified in taking the earth to be stationary? Take the ratio of the radius of the earth to that of the moon to be four.
9. What is meant by impact parameter? What should be the order of magnitude of the impact parameter, if the collision is to be a head-on collision?
10. When does the C.M. system coincide with the lab system in case of two-body collision?
11. Distinguish between elastic and inelastic scattering.
12. Distinguish between differential and total cross-section.
13. What is meant by exoergic and endoergic processes?
14. What is represented by 'differential cross-section'? By 'total cross section'?
15. In what way is the use of C.M. frame advantageous as compared to that of a laboratory frame?

PROBLEMS

1. A moving particle of mass m collides with a stationary particle of mass M. What is the maximum energy that can be transferred? For what value of M is it largest? Hence explain why heavy water is used as a moderator in nuclear reactors.
2. In a head-on collision, a particle moving with velocity V strikes a stationary particle having equal mass. Find the velocity of the two particles after the collision if (a) half the original kinetic energy is lost, (b) the final kinetic energy is 50% greater than the original kinetic energy.
3. Two identical particles, each of mass m and charge e are initially far away from each other. One of the particles is at rest and the other

of the electron and $q = |\mathbf{q}|$ is the magnitude of the phonon wave vector. Also calculate the scattering angle if $E(k) = 2eV$ and the wavelength of the phonon is 10 Å. Assume velocity of sound to be 10^5 cm/s. For this, what is the required angle between **k** and **q** before the collision?

13. The molecules in a gas may be treated as identical hard spheres. Find out how many collisions are required, on an average, to reduce the velocity of an exceptionally fast molecule by a factor of 1000. (Assume that its velocity is so large that the other molecules are effectively at rest even after the molecule has slowed down.) How would the result be affected if the fast molecule has mass m and the others have mass M, the size remaining the same?

14. Consider an endoergic reaction in which the masses do not change Show that, for energies below a certain limit, projectiles with two different energies are observed at the same angle in the lab frame. Use transformation from the C.M. to the lab frame. For a given energy, what is the upper limit on the lab angle? What is the limiting projectile energy below which two energies are observed?

15. At low energies protons and neutrons behave roughly like hard spheres of radius about 2.5×10^{-12} cm. A parallel beam of neutrons, with a flux of 3×10^6 neutrons cm^{-2} s^{-1} strikes a target containing 4×10^{22} protons. A circular detector of radius 2 cm is placed 70 cm away from the target at an angle of 30° to the direction of the beam. Calculate the rate of detection of neutrons and protons.

16. Alpha particles are scattered by a U^{238} nucleus whose radius is $1.5 \times A^{1/3} \times 10^{-13}$ cm; (a) Show that below a limiting energy no α-particle can reach the nucleus. Find this limit; (b) for any given energy above this limit, the α-particles which reach the nucleus are scattered through a certain angle θ_0. Find θ_0 as a function of E; and (c) how can this result be used to find the radius? (Take $M_U/M_\alpha \gg 1$)

17. An atom having velocity V_i emits a photon at angle θ to its direction of motion. Calculate the photon energy after making corrections for recoil. If the mass of the atom is very much greater than $h\nu/c^2$, show that the Doppler effect is observed.

18. A stationary particle of mass $3m$ kg explodes into three pieces of equal mass. Two of the pieces fly off at right angles to each other with speeds $2a$ m/s and $3a$ m/s respectively. What is the magnitude and direction of the momentum of the third fragment? The explosion takes place in 10^{-5} s. Find the average force acting on each piece during the explosion.

19. A particle of mass M_1 and velocity V_1 is captured by a nucleus at rest. A light particle of mass M_2 is ejected at 90° with the path of M_1 with speed V_2, the rest of the nucleus (mass M_3) recoiling with

What is the dependence of the cross-section on the change of energy of the first (or second) type of particles?

28. The average life time of a muon at rest is 2.2×10^{-6} sec. If the muons formed high in the atmosphere travel with speed $V = 0.99c$, what is the average distance they will be observed to traverse before decaying?

29. The maximum energy of electrons produced in the decay of a muon at rest is found to be 55 MeV. The decay is represented by

$$\bar{\mu} \rightarrow \bar{e} + \nu_e + \bar{\nu}_\mu,$$

where ν_e and $\bar{\nu}_\mu$ represent respectively the electron's neutrino and muon's anti-neutrino. Given that the rest mass of the electron is 0.51 MeV and that the neutrino and anti-neutrino have zero rest masses, find the rest mass of the muon. Find the minimum energy carried away by the two neutrinos.

30. A pion whose rest mass is 273 times the rest mass of an electron decays while at rest into a muon of rest mass 207 m_e and a muon's neutrino:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu$$
$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$$

Find the energy and linear momentum received by the muon.

8 Lagrangian Formulation

While solving specific problems we have written the equation of motion of a particle in terms of the cartesian or the polar coordinates. Consider, for example, the motion of a particle in a central force-field. We studied this motion by using plane polar coordinates r and θ. The motion of a projectile is considered in the cartesian coordinate system. It is seen that a particular coordinate system chosen from the symmetry of the physical system helps to simplify the problem. Such a dependence on the coordinate system, however, is undesirable and we should be able to write the equations of motion without any specific reference to the coordinate system used. This is the approach in the Lagrangian formulation of Classical Mechanics. It has, as will be seen in this chapter, a number of advantages over the Newtonian formulation.

This approach may be compared to the use of vectors in describing physical quantities. When expressed through vectors, these equations do not depend on any specific coordinate system and the physical significance of the equation is also clearly seen. This is, however, a very limited comparison. The Lagrangian formulation is of a very general nature and, as we shall see presently, makes use of generalised coordinates and velocities which are independent of the coordinate system and may not be the usual spatial coordinates and velocities.

8.1 CONSTRAINTS

Consider the motion of a free particle. To describe this motion we can use three independent coordinates such as the cartesian coordinates x, y, z or the spherical polar coordinates r, θ, φ and so on. The particle is free to execute motion along any one axis independently with change in one coordinate only. The above statement is equivalent to saying that the particle has three degrees of freedom. The number of independent ways in which a mechanical system can move without violating any constraints

which may be imposed on the system is called the number of degrees of freedom of that system. In other words, the number of degrees of freedom is the number of independent variables which must be simultaneously specified in order to describe the positions and velocities of all particles in the system for any motion which does not violate the constraints. Thus, for a system of N particles moving independently of each other, the number of degrees of freedom is obviously $3N$. For a particle constrained to move on a plane only two variables x, y or r, θ are sufficient to describe its motion and the particle is said to have two degrees of freedom. Thus, the constraint on the motion of the particle in a plane reduces the number of degrees of freedom by one. Further, let the particle be tied to one end of a rigid rod and be capable of rotating about the other end in the plane. The particle is then constrained to move along a circle of radius equal to the length of the rod. The motion of the particle can now be described in terms of a single variable θ and has only one degree of freedom.

When the motion of a system is restricted in some way, constraints are said to have been introduced. A bead sliding down a wire, a disc rolling down an inclined plane, etc. are some illustrations of constrained motion. Thus, when constraints are introduced into a system its number of degrees of freedom is reduced. Very often, we can express constraints in terms of certain equations. Thus, in the case of a rigid body, the constraint that the distance between any two particles of the body is a constant can be written as

$$|\mathbf{r}_i - \mathbf{r}_j|^2 = |\mathbf{r}_{ij}|^2 = (x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2 = \text{const} \tag{8.1}$$

where $\mathbf{r}_i$ and $\mathbf{r}_j$ are the position vectors of the ith and jth particles, respectively.

In the case of a simple pendulum moving in the xy-plane (Fig. 8.1), the two equations of the constraints are

$$z = 0 \text{ and } x^2 + y^2 = l^2 = \text{constant} \tag{8.2}$$

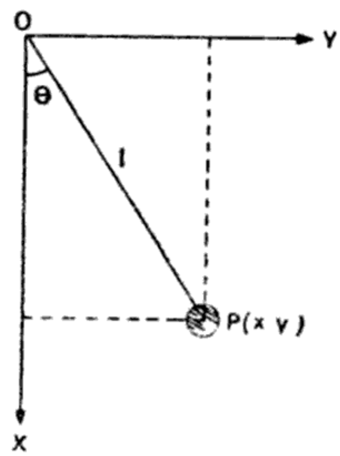


Fig. 8.1 Constraints on a simple pendulum.

It is clear from Fig. 8.1 that only one variable θ is sufficient to locate oscillating particle P

The equation of constraint in the case of a particle moving on or outside the surface of a sphere of radius a is

$$x^2 + y^2 + z^2 \geqslant a^2 \tag{8.3}$$

if the origin of the coordinate system coincides with the centre of the sphere (Fig. 8.2). The equality in expression (8.3) holds as long as the

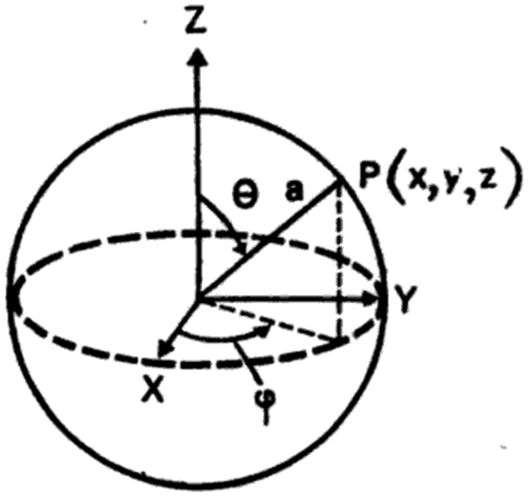


Fig. 8.2 Generalised coordinates of a particle constrained to move on the surface of a sphere

particle is in contact with the surface of the sphere. The inequality in expression (8.3) corresponds to a case when the particle leaves the surface, i.e. $r = \sqrt{x^2 + y^2 + z^2} > a$. Thus, a constraint is a restriction on the freedom of motion of a system of particles in the form of a condition which must be satisfied by their coordinates or by the allowed changes in their coordinates.

(a) Holonomic and Non-Holonomic Constraints

A holonomic constraint is one that may be expressed in the form of an equation relating the coordinates of the system and time. The general form of such equations for a system of N particles is

$$F_i[x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_N, y_N, z_N, t] = 0 \tag{8.4}$$

where $i = 1, 2, 3, \ldots, k$, and F_i is some function of the coordinates. In this equation, i denotes the ith constraint. It should be noted that all variables may not enter in the equation of constraints.

A non-holonomic constraint cannot be expressed in the type or form of equation (8.4). It may be in the form of an inequality as in expression (8.3).

If there are k equations of constraints, the number of degrees of freedom is reduced to $(3N - k)$. Hence, instead of $3N$ coordinates (x_i, y_i, z_i) where $i = 1, 2, \ldots, N$, we can assign the $(3N - k)$ independent varibles $q_1, q_2, \ldots, q_{3N-k}$ to describe the system. Such variables are

denoted by q_i, where $i = 1, 2, \ldots, (3N - k)$. These variables need not necessarily have the dimensions of length or angles, they may be even charges. Such variables are called the generalised coordinates. Thus, $q = \theta$ in the case of a simple pendulum, coordinates $q_1 = r$ and $q_2 = \theta$ in the case of the motion of a particle in a central force field are generalised coordinates. A set of independent coordinates $q_1, q_2, \ldots, q_{3N-k}$ is called a proper set of generalised coordinates.

It should be noted that the constraints on a system are the consequences of the forces exerted on the system by the constraining mechanism. This mechanism may not generally be known. Hence, it is necessary to eliminate the forces of constraint. This is achieved in the Lagrangian formulation by introducing a proper set of generalised coordinates. The choice of the generalised coordinates incorporates the constraints. This is possible only when the constraints are holonomic. We mostly come across systems in which only holonomic constraints have been introduced, hence we shall consider only the holonomic constraints in this book.

The non-holonomic constraints cannot be eliminated by any general method. Each problem involving such constraints needs to be solved individually.

It should be noted that the constraints need not always be written in the form of equations connecting the coordinates, but they can be expressed as well in terms of the velocities. For example, a disc of radius a rolling down an inclined plane along the line of greatest slope will have its constraint equation as follows (see Fig. 8.3). In this case

$$\frac{dl}{dt} = a\dot{\theta}$$

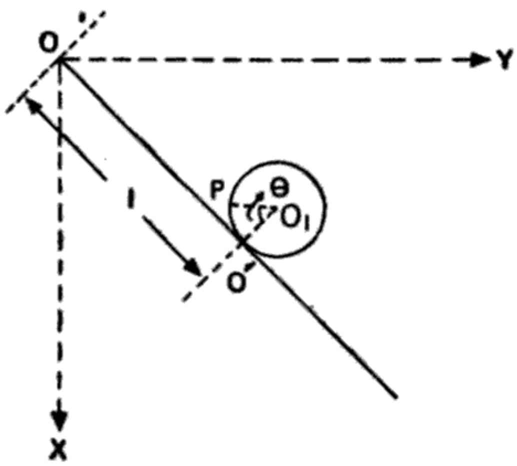


Fig. 8.3 Generalised coordinates of a disc rolling down an inclined plane without slipping

This can also be written as

$$dl = a\, d\theta$$

which after integration reduces to

$$l - a\theta = \text{const}$$

the form

$$\sum (\mathbf{F}_i - \dot{\mathbf{p}}_i)\cdot\delta\mathbf{r}_i = 0 \tag{8.17}$$

Equation (8.17) is the mathematical statement of D'Alembert's principle. In this equation all forces $\mathbf{F}_i$ are the applied forces. The forces of constraints do not appear in this equation for reasons mentioned above.

We now transform D'Alembert's principle into expressions containing independent generalised coordinates only. For this, we consider an infinitesimal virtual displacement $\delta\mathbf{r}_i$ at particular instant t. Then, from equation (8.6), we have

$$\delta\mathbf{r}_i = \sum_j \frac{\partial\mathbf{r}_i}{\partial q_j}\,\delta q_j \tag{8.18}$$

Variation with respect to time is absent in equation (8.18) because virtual displacement $\delta\mathbf{r}_i$ is assumed to take place at fixed instant t. Further the velocities are given by

$$\mathbf{v}_i = \dot{\mathbf{r}}_i = \sum_j \frac{\partial\mathbf{r}_i}{\partial q_j}\,\dot{q}_j + \frac{\partial\mathbf{r}_i}{\partial t} \tag{8.19}$$

where $\dot{q}_j = \dfrac{\partial q_j}{\partial t}$ are the generalised velocities. It should be noted that since generalised coordinates q_j need not have the dimensions of length, generalised velocities $\dot{q}_j$ need not have the dimensions of velocity.

The virtual work done by forces $\mathbf{F}_i$ in terms of virtual displacements $\delta\mathbf{r}_i$ is given by

$$\begin{aligned}\delta W &= \sum_i \mathbf{F}_i\cdot\delta\mathbf{r}_i\\ &= \sum_j\left(\sum_i \mathbf{F}_i\cdot\frac{\partial\mathbf{r}_i}{\partial q_j}\right)\delta q_j\\ &= \sum_j Q_j\,\delta q_j\end{aligned} \tag{8.20}$$

where $Q_j = \sum_i \mathbf{F}_i\cdot\dfrac{\partial\mathbf{r}_i}{\partial q_j}$ represents the jth component of the generalised force. We can express it as

$$\begin{aligned}Q_j &= \sum_i \mathbf{F}_i\cdot\frac{\partial\mathbf{r}_i}{\partial q_j}\\ &= \sum_i\left(F_{ix}\frac{\partial x_i}{\partial q_j} + F_{iy}\frac{\partial y_i}{\partial q_j} + F_{iz}\frac{\partial z_i}{\partial q_j}\right)\end{aligned} \tag{8.21}$$

It is obvious that just as generalised coordinates q_j need not have the dimensions of length, generalised force Q_j need not have the dimensions of force. But product $\sum_j Q_j\,\delta q_j$ must have the dimensions of work.

8.4 LAGRANGE'S EQUATIONS

We have already stated D'Alembert's principle in the form $\sum(\mathbf{F}_i - \dot{\mathbf{p}}_i)\cdot\delta\mathbf{r}_i = 0$ in equation (8.17). The first term, viz. $\sum_i \mathbf{F}_i\cdot\delta\mathbf{r}_i$, has already been proved equal to $\sum_j Q_j\,\delta q_j$ in equation (8.20).

Consider the second term on the right-hand side of this equation. We wish to express it in terms of the virtual displacements of the generalised coordinates. Thus, we can write

$$\sum_i \dot{\mathbf{p}}_i \cdot \delta \mathbf{r}_i = \sum_i \frac{d}{dt}(m_i \dot{\mathbf{r}}_i) \cdot \delta \mathbf{r}_i$$

$$= \sum_{ij} m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j \tag{8.22}$$

The coefficient of δq_j on the right-hand side of equation (8.22) can be written as

$$\sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \sum_i \left[\frac{d}{dt}\left(m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}\right) - m_i \dot{\mathbf{r}}_i \cdot \frac{d}{dt}\left(\frac{\partial \mathbf{r}_i}{\partial q_j}\right)\right] \tag{8.23}$$

Now consider the part $\frac{d}{dt}\left(\frac{\partial \mathbf{r}_i}{\partial q_j}\right)$ of the last term in equation (8.23). In this we can interchange the order of differentiation by making use of equation (8.19). Thus

$$\frac{d}{dt}\left(\frac{\partial \mathbf{r}_i}{\partial q_j}\right) = \sum_k \frac{\partial^2 \mathbf{r}_i}{\partial q_j \, \partial q_k} \dot{q}_k + \frac{\partial^2 \mathbf{r}_i}{\partial q_j \, \partial t}$$

$$= \frac{\partial \dot{\mathbf{r}}_i}{\partial q_j} \tag{8.24}$$

by using equation (8.19) which gives the right-hand side as the expansion of $\frac{\partial \dot{\mathbf{r}}_i}{\partial q_j}$. Thus, equation (8.24) shows that the order of differentiation with respect to t and q_j can be interchanged.

Note that, in equation (8.24), a term dependent upon time is also included. Moreover, from equation (8.19), we have after differentiation with respect to $\dot{q}_j$

$$\frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j} = \frac{\partial \mathbf{r}_i}{\partial q_j} \tag{8.25}$$

Substituting from equations (8.24) and (8.25) into equation (8.23), we get

$$\sum m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \sum_i \left[\frac{d}{dt}\left(m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j}\right) - m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial q_j}\right]$$

$$= \frac{d}{dt}\left[\frac{\partial}{\partial \dot{q}_j}\left(\sum_i \tfrac{1}{2} m_i |\dot{\mathbf{r}}_i|^2\right)\right] - \frac{\partial}{\partial q_j}\left[\left(\sum_i \tfrac{1}{2} m_i |\dot{\mathbf{r}}_i|^2\right)\right]$$

$$= \frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} \tag{8.26}$$

where T, the kinetic energy of the system, is given by

$$T = \sum_i \tfrac{1}{2} m_i |\dot{\mathbf{r}}_i|^2 \tag{8.27}$$

On combining equations (8.26), (8.22) and (8.20), D'Alembert's principle becomes

$$\sum_j \left[\left(\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j}\right) - Q_j\right] \delta q_j = 0 \tag{8.28}$$

Virtual displacements δq_j are all independent of one another. Hence, the set of n equations (8.28), where $j = 1, 2, 3, \ldots, n$, is true only if the

coefficient of each displacement is zero. Thus, we get a set of n equations

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = Q_j \tag{8.29}$$

The equations are valid in the case of conservative as well as non-conservative forces. These equations are called Lagrange's equations. However, this name is used mostly for equations of the systems whenever conservative forces are acting. In that case, the forces can be derivable from potential energy function V. Thus

$$F_{ix} = -\frac{\partial V}{\partial x_i},\ F_{iy} = -\frac{\partial V}{\partial y_i} \text{ and } F_{iz} = -\frac{\partial V}{\partial z_i} \tag{8.30}$$

In the vector notation, we can write

$$\mathbf{F}_i = -\nabla_i V$$

where $$\nabla_i = \mathbf{i}\frac{\partial}{\partial x_i} + \mathbf{j}\frac{\partial}{\partial y_i} + \mathbf{k}\frac{\partial}{\partial z_i}$$

Potential energy function V is a function of $\mathbf{r}_i$ or q_j and is not a function of velocities $\dot{\mathbf{r}}_i$ or $\dot{q}_j$. Under these circumstances, the generalised forces are given by

$$\begin{aligned} Q_j &= \sum_i F_{ix}\frac{\partial x_i}{\partial q_j} + F_{iy}\frac{\partial y_i}{\partial q_j} + F_{iz}\frac{\partial z_i}{\partial q_j} \\ &= -\sum_i \left[\frac{\partial V}{\partial x_i}\frac{\partial x_i}{\partial q_j} + \frac{\partial V}{\partial y_i}\frac{\partial y_i}{\partial q_j} + \frac{\partial V}{\partial z_i}\frac{\partial z_i}{\partial q_j}\right] \\ &= -\frac{\partial V}{\partial q_j} \end{aligned} \tag{8.31}$$

Thus, the relation between potential energy function V and component Q_j of generalised conservative force is of the same form as given in equation (8.30). Moreover, $\frac{\partial V}{\partial \dot{q}_j} = 0$.

Hence, equation (8.29) can be written as

$$\frac{d}{dt}\frac{\partial (T-V)}{\partial \dot{q}_j} - \frac{\partial (T-V)}{\partial q_j} = 0 \tag{8.32}$$

Let the Lagrangian function L be defined by

$$\begin{aligned} L &= L(q_1, q_2, q_3, \ldots, q_n; \dot{q}_1, \dot{q}_2, \dot{q}_3, \ldots, \dot{q}_n) \\ &= T - V \end{aligned} \tag{8.33}$$

Then, equations (8.32) become

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0 \tag{8.34}$$

which are known as Lagrange's equations of motion for a conservative holonomic system.

We have derived Lagrange's equations from D'Alembert's principle in which we have explicitly used Newton's equations of motion. Thus, the derivation of Lagrange's equations for a system is equivalent to Newton's equation of motion.

the velocity. Squaring equation (8.36), we have

$$\dot{r}_i^2 = \sum_{jk} \frac{\partial \mathbf{r}_i}{\partial q_j} \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} \dot{q}_j \dot{q}_k + 2 \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} \cdot \frac{\partial \mathbf{r}_i}{\partial t} \dot{q}_j + \left(\frac{\partial \mathbf{r}_i}{\partial t}\right)^2 \tag{8.37}$$

Substituting this value of $\dot{r}_i^2$ in equation (8.35), we get

$$T = \sum_{jk} a_{jk} \dot{q}_j \dot{q}_k + \sum_j b_j \dot{q}_j + c \tag{8.38}$$

where

$$\left.\begin{aligned} a_{jk} &= \tfrac{1}{2} \sum_i m_i \frac{\partial \mathbf{r}_i}{\partial q_j} \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} \\ b_j &= \sum_i m_i \frac{\partial \mathbf{r}_i}{\partial q_j} \cdot \frac{\partial \mathbf{r}_i}{\partial t} \\ \text{and} \qquad c &= \tfrac{1}{2} \sum_i m_i \left(\frac{\partial \mathbf{r}_i}{\partial t}\right)^2 \end{aligned}\right\} \tag{8.39}$$

The case when transformation equations (8.6) are independent of time is of particular interest. The constraints are then scleronomous and the partial derivatives with respect to time vanish. Hence, $b_j = 0$ and $c = 0$. In this case, the kinetic energy of the system is a homogeneous quadratic function of the generalised velocities. Thus

$$T = \sum_{jk} a_{jk} \dot{q}_j \dot{q}_k \tag{8.40}$$

Let us now form the sum $\sum_l \dot{q}_l \frac{\partial T}{\partial \dot{q}_l}$. From equation (8.40), we have

$$\frac{\partial T}{\partial \dot{q}_l} = \sum_i a_{il} \dot{q}_i + \sum_j a_{lj} \dot{q}_j$$

While carrying out the differentiation, we have used $\frac{\partial \dot{q}_i}{\partial \dot{q}_l} = \delta_{il}$, since $\dot{q}_j$ are a set of independent velocities. Here, the suffixes i and j are dummy indices as these can be replaced by other suffixes without changing the value of the sum. The sum $\sum \dot{q}_l \frac{\partial T}{\partial \dot{q}_l}$ is then written as

$$\sum_l \dot{q}_l \frac{\partial T}{\partial \dot{q}_l} = \sum_{il} a_{il} \dot{q}_i \dot{q}_l + \sum_{jl} a_{lj} \dot{q}_l \dot{q}_j$$

Since, all the suffixes are dummy and are symmetrical in i and j, the two sums on the right-hand side are identical. Hence

$$\sum_l \dot{q}_l \frac{\partial T}{\partial \dot{q}_l} = 2 \sum_{il} a_{il} \dot{q}_i \dot{q}_l = 2T \tag{8.41}$$

The result of the equation follows immediately from Euler's theorem which states that if $f(x_i)$ is a homogeneous function of the nth degree of a set of variables x_i, then

$$\sum x_i \frac{\partial f}{\partial x_i} = nf \tag{8.42}$$

Thus, in obtaining Lagrange's equation of motion of a system, we have to form the Lagrangian L of that system, which is the difference between the kinetic and potential energy of the system. Once this is done, the

remaining procedure is mechanical. But, it is very important to get a correct expression for kinetic energy T. It is always better for a beginner to start from the expression for T in the cartesian coordinates and then substitute proper values of $\dot{x}_i$, $\dot{y}_i$ and $\dot{z}_i$ in it from the transformation equations.

The procedure is generally lengthy and tedious, and it is sometimes found convenient to write the velocity directly in terms of the generalised coordinates and obtain the expression for the kinetic energy.

Consider an illustration of a double pendulum which is a system of two pendulums, the second pendulum of length l_2 being suspended from the bob of the first pendulum of length l_1 (Fig. 8.4). Let us assume that the pendulums move in the xy plane only. The generalised coordinates in this case are $q_1 = \theta_1$ and $q_2 = \theta_2$ which are the angular displacements of the first and second pendulums from the vertical respectively. It is difficult to find an expression for kinetic energy T directly in terms of $\dot{\theta}_1$ and $\dot{\theta}_2$. We, therefore, write the transformation equations of the two bobs whose coordinates are (x_1, y_1) and (x_2, y_2) respectively. From Fig. 8.4, we have

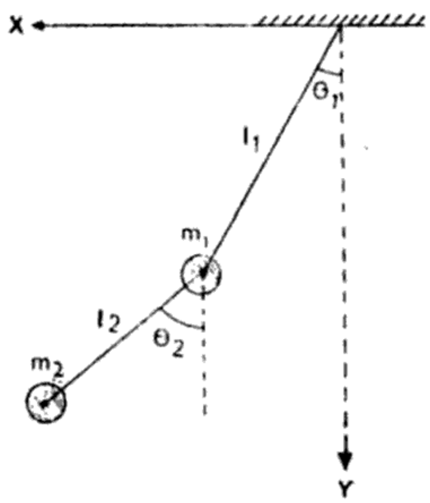


Fig. 8.4 Double pendulum

$$\left.\begin{aligned} x_1 &= l_1 \sin\theta_1 \\ y_1 &= l_1 \cos\theta_1 \\ x_2 &= l_1 \sin\theta_1 + l_2 \sin\theta_2 \\ \text{and} \qquad y_2 &= l_1 \cos\theta_1 + l_2 \cos\theta_2 \end{aligned}\right\} \tag{8.43}$$

Differentiating equations (8.43) with respect to time, we get

$$\left.\begin{aligned} \dot{x}_1 &= l_1 \cos\theta_1\dot{\theta}_1 \\ \dot{y}_1 &= -l_1 \sin\theta_1\dot{\theta}_1 \\ \dot{x}_2 &= l_1 \cos\theta_1\dot{\theta}_1 + l_2 \cos\theta_2\dot{\theta}_2 \\ \text{and} \qquad \dot{y}_2 &= -l_1 \sin\theta_1\dot{\theta}_1 - l_2 \cos\theta_2\dot{\theta}_2 \end{aligned}\right\} \tag{8.44}$$

Now, kinetic energy T is given by

$$T = \tfrac{1}{2}m_1(\dot{x}_1^2 + \dot{y}_1^2) + \tfrac{1}{2}m_2(\dot{x}_2^2 + \dot{y}_2^2)$$

Substituting the values of $\dot{x}_1, \dot{y}_1, \ldots,$ etc. and simplifying, we get

$$T = \tfrac{1}{2}(m_1 + m_2)l_1^2\dot{\theta}_1^2 + \tfrac{1}{2}m_2 l_2^2\dot{\theta}_2^2 + m_2 l_1 l_2 \cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 \tag{8.45}$$

This expression can also be obtained by using equation (8.40). Thus

$$T = \sum_{jk} a_{jk}\dot{q}_j\dot{q}_k$$

$$= a_{\theta_1\theta_1}\dot{\theta}_1^2 + (a_{\theta_1\theta_2} + a_{\theta_2\theta_1})\dot{\theta}_1\dot{\theta}_2 + a_{\theta_2\theta_2}\dot{\theta}_2^2 \tag{8.46}$$

where

$$a_{\theta_1\theta_1} = \tfrac{1}{2}m_1\left[\left(\frac{\partial x_1}{\partial \theta_1}\right)^2 + \left(\frac{\partial y_1}{\partial \theta_1}\right)^2\right] + \tfrac{1}{2}m_2\left[\left(\frac{\partial x_2}{\partial \theta_1}\right)^2 + \left(\frac{\partial y_2}{\partial \theta_1}\right)^2\right]$$

$$a_{\theta_1\theta_2} = a_{\theta_2\theta_1} = \tfrac{1}{2}m_1\left[\frac{\partial x_1}{\partial \theta_1}\frac{\partial x_1}{\partial \theta_2} + \frac{\partial y_1}{\partial \theta_1}\frac{\partial y_1}{\partial \theta_2}\right] + \tfrac{1}{2}m_2\left[\frac{\partial x_2}{\partial \theta_1}\frac{\partial x_2}{\partial \theta_2} + \frac{\partial y_2}{\partial \theta_1}\frac{\partial y_2}{\partial \theta_2}\right]$$

and $$a_{\theta_2\theta_2} = \tfrac{1}{2}m_1\left[\left(\frac{\partial x_1}{\partial \theta_2}\right)^2 + \left(\frac{\partial y_1}{\partial \theta_2}\right)^2\right] + \tfrac{1}{2}m_2\left[\left(\frac{\partial x_2}{\partial \theta_2}\right)^2 + \left(\frac{\partial y_2}{\partial \theta_2}\right)^2\right]$$

Substituting the values of the various derivatives and simplifying, we get

$$a_{\theta_1\theta_1} = \tfrac{1}{2}(m_1 + m_2)l_1^2$$

$$a_{\theta_1\theta_2} = a_{\theta_2\theta_1} = \tfrac{1}{2}m_2 l_1 l_2 \cos(\theta_1 - \theta_2) \tag{8.47}$$

and $$a_{\theta_2\theta_2} = \tfrac{1}{2}m_2 l_2^2$$

Substituting these values of a's in the expression for T, we get the same expression for the kinetic energy as before.

8.6 SYMMETRIES AND THE LAWS OF CONSERVATION

So far, we have discussed Lagrange's dynamical equations of a system. If the system under consideration has n degrees of freedom, we get n second-order differential equations. The solution of each equation will need the evaluation of a double integral and will, therefore, involve two constants—initial position q_0 and initial velocity $\dot{q}_0$. Naturally, the solutions of n differential equations will involve $2n$ constants—n values of initial q's and $\dot{q}$'s.

In many problems, the solution cannot be obtained in terms of known functions. Moreover, sometimes solutions of the type $q_j = q_j(t)$ are of no interest to us. For example, when we study the motion of systems consisting of atoms and molecules, we are only interested in the evaluation of quantities such as energies and angular momenta. However, information regarding the physical nature of the motion of the system can often be extracted without integrating the equations of motion.

On considering the symmetries of the system, one can immediately obtain first integrals of the equations of motion. The first integrals are constants of motion. These are the first-order differential equations of the type

$$f(q_1, q_2, \ldots, q_n; \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n, t) = \text{const} \tag{8.48}$$

It is obvious that the first integral contains the first derivative of q's. The first integrals reveal a lot of information regarding the system under consideration. In fact, the conserved quantities discussed in Chapter 3 are first integrals of motion.

Consider a system of particles in a conservative force-field. Then, potential energy V depends only upon the position and we have

$$\frac{\partial L}{\partial \dot{x}_i} = \frac{\partial (T - V)}{\partial \dot{x}_i} = \frac{\partial}{\partial \dot{x}_i}\sum_j \tfrac{1}{2}m_j(\dot{x}_j^2 + \dot{y}_j^2 + \dot{z}_j^2)$$

$$= m_i \dot{x}_i = p_{x_i}$$

where p_{x_i} is the x-component of the linear momentum of the ith particle of the system. We can generalise this result and define the generalised momentum by the formula

$$p_j = \frac{\partial L}{\partial \dot{q}_j} \tag{8.49}$$

Quantity p_j is very often called the *canonical* or the *conjugate* momentum. If q_j represents translation (i.e. linear displacement), then p_j represents linear momentum (i.e., mass multiplied by velocity). But, if q_j represents an angle, p_j represents the angular momentum. The generalised definition of momentum allows us to consider non-mechanical systems as well. If we consider a charged particle moving in an electromagnetic field, then the Lagrangian is defined as

$$L = \tfrac{1}{2}m|\dot{\mathbf{r}}|^2 - q\Phi + q\mathbf{A}\cdot\dot{\mathbf{r}}$$

where Φ is a scalar potential and $\mathbf{A}$ is a vector potential. Differentiating with respect to $\dot{x}$, we get

$$p_x = m\dot{x} + qA_x$$

as the canonical momentum to x. It is not the usual kinetic momentum $m\dot{x}$ but has a contribution of qA_x from the electromagnetic field.

It should be noted that with this definition, Lagrange's equations for a conservative system assume a simple form similar to Newton's equations of motion. These are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} = \dot{p}_j = \frac{\partial L}{\partial q_j} = -\frac{\partial V}{\partial q_j} = Q_j$$

or

$$\dot{p}_j = Q_j \tag{8.50}$$

8.7 CYCLIC OR IGNORABLE COORDINATES

The Lagrangian L is written as a function of q_j and $\dot{q}_j$. If any one coordinate, say q_k, is absent in the expression for the Lagrangian L, then the partial derivative of L with respect to q_k will vanish, i.e.

$$\frac{\partial L}{\partial q_k} = 0$$

Thus, the equation of motion corresponding to variable q_k becomes

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) = 0 \tag{8.51}$$

Integrating equation (8.51), we get

$$\frac{\partial L}{\partial \dot{q}_k} = p_k = \text{const}$$

Thus, whenever coordinate q_k does not appear explicitly in the Lagrangian function L, corresponding linear momentum p_k is a constant of the motion, or in other words, it is the first integral of the motion. Such a coordinate q_k is said to be *cyclic* or *ignorable.*

Thus, we are led to the conservation of momentum—linear or angular—conjugate to the cyclic coordinate. The laws of conservation discussed in

Chapter 3 are, in general, contained in the conservation theorem relating to the cyclic coordinates.

There should be a deeper relationship between a cyclic coordinate and the physical nature of the motion of the system. When generalised translation coordinate q_k is cyclic, i.e., absent in the Lagrangian function, it means that the system can be translated along q_k without any effect on the Lagrangian. The Lagrangian function L is unchanged under this translation and the equation of motion remains the same. Thus, the system is invariant under translation or the system is said to have translational symmetry. This immediately leads us to the conservation of conjugate momentum. To illustrate this, consider the motion of two bodies under the action of internal forces. As no external force is acting on the system, the translation of the whole system will not affect the internal forces or the velocities. Thus, kinetic energy T, potential energy V and hence, the Lagrangian function L are unaffected. If $\mathbf{R}(X, Y, Z)$ represents the position vector of the centre of mass of the system, then

$$\frac{\partial L}{\partial X} = \frac{\partial L}{\partial Y} = \frac{\partial L}{\partial Z} = 0$$

Hence, the linear momentum of the centre of mass of the system, i.e., $M\dot{\mathbf{R}}$ is conserved. This result was obtained in Chapter 3. Similar considerations can be applied to a system having rotational symmetry. If the rotation of the system about a particular axis leaves the system unaffected, then the corresponding coordinate is cyclic and the component of the angular momentum in that direction is conserved.

We have seen that the conservation of the linear and the angular momentum of a system corresponds to its translational and rotational symmetries. This approach involving the considerations of symmetries is very fundamental and is helpful in many modern fields of physics.

It, therefore, appears that the conservation of energy of a system must also correspond to some symmetry of the system. This symmetry is with respect to the translation of the system along time. If the Lagrangian function L of a system does not explicitly depend upon time, then the translation of the system along time t leaves it unchanged. Then, we can write

$$\frac{\partial L}{\partial t} = 0 \tag{8.52}$$

Hence, the total time derivative of $L \equiv L(q_j, \dot{q}_j)$ is given by

$$\frac{dL}{dt} = \sum_j \frac{\partial L}{\partial q_j} \dot{q}_j + \sum_j \frac{\partial L}{\partial \dot{q}_j} \ddot{q}_j \tag{8.53}$$

It should be noted that the partial time derivative does not occur in equation (8.53) where we have used equation (8.52). Using Lagrange's equation, we can rewrite equation (8.53) as

$$\frac{dL}{dt} = \sum_j \dot{q}_j \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) + \sum_j \frac{\partial L}{\partial \dot{q}_j} \ddot{q}_j$$

Hence,
$$\frac{dL}{dt} = \sum_j \frac{d}{dt}\left(\dot{q}_j \frac{\partial L}{\partial \dot{q}_j}\right)$$

or
$$\frac{d}{dt}\left(L - \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j}\right) = 0 \tag{8.54}$$

Thus, the quantity in the bracket on the left-hand side of equation (8.54) is constant in time. It is denoted by $-H$. Hence, we have

$$H = \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L$$
$$= \sum_j \dot{q}_j p_j - L \tag{8.55}$$

If, potential energy V does not depend upon velocity or time, i.e. $V \equiv V(q_j)$ only, we can write

$$p_j = \frac{\partial L}{\partial \dot{q}_j} = \frac{\partial}{\partial \dot{q}_j}(T - V) = \frac{\partial T}{\partial \dot{q}_j}, \text{ simply.}$$

With this, the first term in equation (8.55) is $2T$ from equation (8.41). The transformation equations connecting $\mathbf{r}_i$ and q_j do not contain time explicitly, i.e., the constraints and the forces are independent of time and the kinetic energy is a quadratic function of velocities. Hence, equation (8.55) becomes

$$H = 2T - L = 2T - (T - V) = T + V = E \tag{8.56}$$

where E is the total energy of the system and is a constant of the motion.

Function H introduced in equation (8.55) is called the Hamiltonian of the system. It is equal to the total energy of the system only when the kinetic energy of the system is a homogeneous quadratic function of the velocities and the potential energy is independent of the velocities and time. The statements that kinetic energy T is a homogeneous quadratic function of velocities, or that the transformation equation $\mathbf{r}_i = \mathbf{r}_i(q_j)$ and the constraints are independent of time, are equivalent.

Thus, we have obtained the law of conservation of energy for a conservative system having time-independent constraints. This last restriction was not necessary for the law in Chapter 3.

In Chapter 3, the changes in energy (kinetic and potential) were due to the work done by all the forces including the forces of constraint. In the Lagrangian formulation, the forces of constraint were removed and the potential energy includes the work done by the applied forces only. When the constraints depend upon time, they also do work during the displacement which is included in potential energy V and hence total energy E will not be equal to the Hamiltonian H. When the constraints are time-independent, it will be recalled that the virtual work done by the forces of constraint is zero. Further, in this case, the virtual displacement which was taken at fixed instant t can now be taken in an interval of time δt and the difference between the real and the virtual displacements disappears. In that case, total energy E will be equal to the Hamiltonian H.

It should be noted that if the constraints are time dependent or the transformation equations $\mathbf{r}_i = \mathbf{r}_i(q_j, t)$ contain time explicitly, $H \neq E$; but still total energy E is conserved.

1. *Illustrations*: Consider the motion of a particle due to force $\mathbf{F}$ having components F_x, F_y and F_z along the three axes of the cartesian coordinates. Kinetic energy T of the particle is given by

$$T = \tfrac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

Therefore

$$\frac{\partial T}{\partial x} = \frac{\partial T}{\partial y} = \frac{\partial T}{\partial z} = 0$$

and
$$\frac{\partial T}{\partial \dot{x}} = m\dot{x},\ \frac{\partial T}{\partial \dot{y}} = m\dot{y} \text{ and } \frac{\partial T}{\partial \dot{z}} = m\dot{z}$$

Hence, the equations of motion are

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{x}} = \frac{d}{dt}(m\dot{x}) = F_x \ldots \text{etc.}$$

These are nothing but Newton's equations of motion.

2. Consider now a bead sliding along a uniformly rotating wire in a force-free space. The transformation equations relating the cartesian and polar coordinates of the bead are

$$x = r\cos\theta = r\cos\omega t$$

and
$$y = r\sin\theta = r\sin\omega t$$

where ω is the constant angular velocity of the wire.

Kinetic energy T is given by

$$\begin{aligned} T &= \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) \\ &= \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) \\ &= \tfrac{1}{2}m(\dot{r}^2 + r^2\omega^2) \end{aligned}$$

Thus, we find that T is not a homogeneous quadratic function of velocities only. The equation of motion can be written as

$$m\ddot{r} = mr\omega^2$$

or
$$\ddot{r} = r\omega^2$$

This is the familiar expression of the centripetal acceleration. This method used to obtain the equation of motion does not at all involve the force of constraint that keeps the bead on the wire.

3. *Atwood's machine*: Atwood's machine (Fig. 8.5) is an illustration of a simple mechanical system with a holonomic constraint. It consists of two masses m_1 and m_2 tied together by means of a light inextensible cord of length l. The cord passes round a light frictionless pulley and the two masses hang on the two sides of the pulley. We find from Fig. 8.5 that only one variable x is independent, since length l of the

The expression for the total energy can be rewritten after substituting for φ as

$$E = \tfrac{1}{2}ml^2\dot{\theta}^2 + \frac{p_\varphi^2}{2ml^2 \sin^2 \theta} + mgl \cos \theta$$
$$= T + V_e$$

where $V_e = mgl \cos \theta + \dfrac{p_\varphi^2}{2ml^2 \sin^2 \theta}$ is the effective potential energy and it depends only upon θ. In case the pendulum is restricted to move only in one plane, say the $\varphi = 0$ plane, the equation of motion reduces to

$$\ddot{\theta} = \frac{g}{l} \sin \theta$$

which is the familiar equation occurring in the case of a simple pendulum with the difference that θ here is measured with respect to the vertically upward direction.

8.8 VELOCITY-DEPENDENT POTENTIAL OF ELECTROMAGNETIC FIELD

We can obtain Lagrange's equation in the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0$$

even if the system is not conservative in the usual sense provided that the generalised forces are expressible in the form

$$Q_j = -\frac{\partial U}{\partial q_j} + \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_j}\right) \tag{8.57}$$

where $U(q_j, \dot{q}_j)$ may be called the 'generalised potential' or the 'velocity-dependent potential'. We come across such a potential when we consider the electromagnetic forces acting on moving charges.

Maxwell's equations are stated as

$$\left.\begin{aligned} \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0, \qquad \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \\ \nabla \times \mathbf{B} &= \mu_0 j + \mu_0\epsilon_0 \frac{\partial \mathbf{E}}{\partial t}, \qquad \nabla \cdot \mathbf{B} = 0 \end{aligned}\right\} \tag{8.58}$$

Quantity $\epsilon_0\mathbf{E} = \mathbf{D}$ is called the electric displacement while the magnetic induction is $\mathbf{B} = \mu_0\mathbf{H}$.

The force on a charged particle having charge q and moving in an electromagnetic field is given by

$$\mathbf{F} = q[\mathbf{E} + \mathbf{v} \times \mathbf{B}] \tag{8.59}$$

and is known as the 'Lorentz force'.

From equation (8.58), we observe that $\nabla \times \mathbf{E} \neq 0$. Hence, we cannot express $\mathbf{E}$ as a gradient of a scalar function. But, $\nabla \cdot \mathbf{B} = 0$ suggests that we can express $\mathbf{B}$ as a curl of vector $\mathbf{A}$. Thus

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{8.60}$$

where **A** is called a magnetic vector potential. With this substitution for **B**, the equation

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0$$

becomes
$$\nabla \times \mathbf{E} + \frac{\partial}{\partial t}(\nabla \times \mathbf{A}) = 0$$

or
$$\nabla \times \left[\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}\right] = 0 \tag{8.61}$$

Hence, we can write

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla \Phi$$

or
$$\mathbf{E} = -\nabla \Phi - \frac{\partial \mathbf{A}}{\partial t} \tag{8.62}$$

The Lorentz force can now be written as

$$\mathbf{F} = q\left[-\nabla \Phi - \frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \times \nabla \times \mathbf{A}\right] \tag{8.63a}$$

Consider the x-components of the various terms of equation (8.63a). We have

$$[\nabla \Phi]_x = \frac{\partial \Phi}{\partial x}$$

$$[\mathbf{v} \times \nabla \times \mathbf{A}]_x = v_y\left[\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right] - v_z\left[\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right] \tag{8.63b}$$

$$= v_y \frac{\partial A_y}{\partial x} + v_z \frac{\partial A_z}{\partial x} - v_y \frac{\partial A_x}{\partial y} - v_z \frac{\partial A_x}{\partial z}$$

Now, we add and subtract term $v_x \dfrac{\partial A_x}{\partial x}$ to the right-hand side of this expression. Then, we get

$$[\mathbf{v} \times \nabla \times \mathbf{A}]_x = v_x \frac{\partial A_x}{\partial x} + v_y \frac{\partial A_y}{\partial x} + v_z \frac{\partial A_z}{\partial x} - v_x \frac{\partial A_x}{\partial x} - v_y \frac{\partial A_x}{\partial y} - v_z \frac{\partial A_x}{\partial z} \tag{8.64}$$

Further, the total time derivative of $A_x \equiv A_x(x, y, z, t)$ is

$$\frac{dA_x}{dt} = \frac{\partial A_x}{\partial t} + \left(\frac{\partial A_x}{\partial x}\frac{dx}{dt} + \frac{\partial A_y}{\partial y}\frac{dy}{dt} + \frac{\partial A_z}{\partial z}\frac{dz}{dt}\right)$$

or
$$\frac{dA_x}{dt} = \frac{\partial A_x}{\partial t} + \left(v_x \frac{\partial A_x}{\partial x} + v_y \frac{\partial A_y}{\partial y} + v_z \frac{\partial A_z}{\partial z}\right) \tag{8.65}$$

Hence, we can write

$$[\mathbf{v} \times \nabla \times \mathbf{A}]_x = \frac{\partial}{\partial x}(\mathbf{v} \cdot \mathbf{A}) - \frac{dA_x}{dt} + \frac{\partial A_x}{\partial t}$$

With these substitutions, the x-component of the Lorentz force is given by

$$F_x = q\left[-\frac{\partial}{\partial x}\{\Phi - \mathbf{v} \cdot \mathbf{A}\} - \frac{d}{dt}\left\{\frac{\partial}{\partial v_x}(\mathbf{A} \cdot \mathbf{v})\right\}\right] \tag{8.66}$$

Now, scalar potential Φ is independent of the velocity; hence we can say that

$$F_x = -\frac{\partial U}{\partial x} + \frac{d}{dt}\frac{\partial U}{\partial v_x}$$

in comparison with equation (8.66), where the generalised potential energy is

$$U = q\Phi - q\mathbf{A}\cdot\mathbf{v} \tag{8.67}$$

With this notation, the Lagrangian for a charged particle moving in an electromagnetic field is given by

$$L = T - U = T - q\Phi + q\mathbf{A}\cdot\mathbf{v} \tag{8.68}$$

This was the Lagrangian used earlier in article 8.6 to derive the generalised momentum of a particle in the electromagnetic field.

8.9 RAYLEIGH'S DISSIPATION FUNCTION

Another point regarding Lagrange's equation must be noted. Only if some of the forces acting on the system are derivable from a potential, Lagrange's equations assume the form

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = Q_j \tag{8.69}$$

where L contains only those forces that are conservative while Q_j includes the forces that are not derivable from a potential. An illustration of this latter type of force is the frictional force. Many times we come across a situation when the frictional force is proportional to velocity. The x-component of this force may be written as

$$F_{fx} = -k_x v_x \tag{8.70}$$

where k_x is the x-component of the frictional force per unit velocity in the x-direction. Forces of this type are derivable from Rayleigh's dissipation function $\mathcal{F}$ defined by

$$\mathcal{F} = \tfrac{1}{2}\sum_i (k_x v_{i_x}^2 + k_y v_{i_y}^2 + k_z v_{i_z}^2) \tag{8.71}$$

where $i = 1, 2, \ldots, n$ covers all the particles of the system.
It is obvious that

$$F_f = -\nabla_v \mathcal{F} \tag{8.72}$$

where we have introduced a new symbol

$$\nabla_v = \mathbf{i}\,\frac{\partial}{\partial v_x} + \mathbf{j}\,\frac{\partial}{\partial v_y} + \mathbf{k}\,\frac{\partial}{\partial v_z}$$

which is the vector velocity differential operator.

To explain the physical significance of Rayleigh's dissipation function, let us calculate the work done by the system against friction as

$$\begin{aligned} dW_f &= -\mathbf{F}_f\cdot d\mathbf{r} = -\mathbf{F}_f\cdot\mathbf{v}\,dt \\ &= (k_x v_x^2 + k_y v_y^2 + k_z v_z^2)\,dt \end{aligned}$$

Therefore, $$\frac{dW_f}{dt} = (k_x v_x^2 + k_y v_y^2 + k_z v_z^2) = 2\mathcal{F}$$

Thus, the rate of dissipation of energy by friction is equal to twice Rayleigh's dissipation function.

Component Q_j of the generalised force arising as a result of frictional

force is given by

$$Q_j = \sum_i \mathbf{F}_{fi} \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = -\sum \nabla_v \mathcal{F} \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}$$

$$= -\sum \nabla_v \mathcal{F} \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j}, \text{ by equation (8.25)}$$

$$= -\frac{\partial \mathcal{F}}{\partial \dot{q}_j}$$

Substituting this value of Q_j, we can write Lagrange's equations as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} + \frac{\partial \mathcal{F}}{\partial \dot{q}_j} = 0 \tag{8.73}$$

Thus, if frictional forces are acting on the system, we must specify two scalar functions—the Lagrangian L and Rayleigh's dissipation function $\mathcal{F}$—to derive the equations of motion.

QUESTIONS

1. What are constraints? Explain, giving examples, the meaning of holonomic and nonholonomic constraints.
2. Explain the meaning of scleronomous and rheonomous constraints. Give illustrations of each.
3. Is the Lagrangian formulation more advantageous than the Newtonian formulation? Why?
4. What do you understand by cyclic coordinates? Show that the generalised momentum corresponding to a cyclic coordinate is a constant of motion.
5. Explain the term 'virtual displacement' and state the principle of virtual work.
6. When is a generalised force the same as the corresponding component of conventional force? When is it different?
7. Describe the use of Rayleigh's dissipation function.
8. What arbitrariness exists in the definition of the Lagrangian?
9. How is a generalised potential defined? What advantage does it have over the conventional potential?
10. What do you understand by generalised momentum? Express the momentum conservation theorem in terms of generalised momenta.
11. State D'Alembert's principle in words.
12. Define the Hamiltonian. When is it equal to the total energy of the system? When is it conserved?
13. What is meant by velocity-dependent potential? Where do we come across such a potential?
14. What is meant by a configuration space? How is this concept used

to describe the motion of a system of particles?

15. Generalised coordinates may not have the dimensions of distance, generalised force may not have the dimensions of force, still their product has the dimensions of work. Explain.
16. While solving a problem in mechanics, a lot of labour is saved by choosing the appropriate coordinate system. The choice is usually guided by the symmetry (spherical, cylindrical, etc.) possessed by the system. In the following cases find out the most convenient coordinate system (try cartesian, spherical polar and cylindrical systems). Further find out the degrees of freedom and all the equations of constraint.
    (i) a double pendulum
    (ii) a disc rolling down an inclined plane
    (iii) a particle constrained to move on the surface of a sphere
    (iv) a particle moving along the surface of a right circular cone
    (v) a missile on a submarine
    (vi) a dumb bell whose centre is constrained to move along a circle
17. Is it possible to express all the *fundamental* forces in nature in the form of generalised potentials? Explain.
18. What is meant by a non-conservative force? Give some examples of such forces. Is the Lorentz force non-conservative? Explain.
19. Conditions for the conservation of generalised momenta corresponding to cyclic coordinates are more general than the two momentum conservation theorems. Comment.

PROBLEMS

1. Write down the convenient generalised coordinates for: (a) a missile in a submarine, free to point in any direction, (b) a point on the wheel of the bicycle, (c) a double pendulum, (d) a spherical pendulum. Also write down the equations of constraints.
2. Write down the Lagrangian in terms of convenient coordinates for: (a) a particle of mass m and charge q, in a uniform gravitational field and the electric field of an infinite vertical line charge of density σ, (b) a planet in the field of the sun and rotating about a fixed axis, (c) two particles of masses m_1 and m_2 and charges q and $-q$ respectively, in a uniform gravitational field. Also obtain the equations of constraints in each case.
3. A cylinder, initially at rest, rolls down an inclined plane without slipping. Investigate its motion using Lagrange's equations. What ensures rolling without slipping? Can we still use Lagrange's equations? Explain.

4. A Kater's pendulum has a radius of gyration k and mass M and its centre of mass is at distance L below the point of support. A pin is attached to it at distance d above the support and from this hangs a simple pendulum of mass m and length l. Obtain the equations of motion of the simple pendulum. Solve these for $M \gg m$. Explain why we get the unphysical result that the simple pendulum has infinite amplitude at resonance.
5. A spring of force constant k is confined to move along a vertical line with its upper end fixed. A simple pendulum of length l and mass m is suspended from its lower end. Obtain the Lagrangian equations and show that they are equivalent to Newton's equations.
6. Show, from Lagrange's equations, that the orbit of a particle in a central force field lies in a plane.
7. Two particles of masses m_1 and m_2 lie on a smooth table connected by a spring of unstretched length l, force constant k and mass m. Obtain the constants of motion and frequency of vibration in the absence of rotation.
8. A simple pendulum of mass m has support of mass M. The support is free to slide on a horizontal frictionless plane. Obtain and solve the equations of motion.
9. A double pendulum vibrates in a vertical plane. If masses of the bobs are m_1 and m_2 and the lengths of the pendulums are l_1 and l_2 respectively, obtain the Lagrangian and Lagrange's equations for this double pendulum
10. Two particles of masses m_1 and m_2 are connected by a string which passes through a hole in a smooth table so that m_1 rests on the table surface and m_2 hangs underneath. Assuming that m_2 moves only in a vertical line, find the suitable generalised coordinates for the system. Write down Lagrange's equations for the system and discuss their physical significance, if any. Reduce the problem to a single second-order differential equation and obtain a first integral of the equation. What is its physical significance? (Consider the motion only so long as neither m_1 nor m_2 passes through the hole.)
11. A hoop rolls without slipping along a horizontal circle and is free to tilt. Obtain its equations of motion. Derive the condition for rolling at constant inclination. Find the horizontal and vertical components of the force at the point of contact.
12. A body is sliding down an inclined plane which is moving horizontally with constant velocity. Find the position of the body as the function of time.
13. Show that for a hoop on a plane, the constraint of rolling without slipping is not holonomic. Write down the constraint in differential form.
14. Show that the kinetic energy of a hoop rolling without slipping on a

23. Viscous force F retarding a sphere in a fluid of coefficient of viscosity η is given by Stokes's law,

$$F = -6\pi\eta av$$

where v is the terminal velocity of the sphere. Set up the corresponding dissipation function. Then set up Lagrange's equation for the sphere falling vertically under the action of gravity in the fluid. Solve it for terminal velocity.

24. An inductor and a capacitor are connected in parallel in a circuit containing an alternating e.m.f. Obtain the Lagrangian function for this system.

25. A particle of mass m moves without friction on a straight wire inclined at angle θ with the vertical z-axis. If the wire rotates about the z-axis with constant angular velocity ω, find the Lagrange's equation of motion and its solution.

9 Moving Coordinate Systems

Very often we come across problems in physics, the solutions of which can be readily obtained if we employ moving coordinate systems or frames of reference. In such problems, we usually consider two frames, one fixed in the laboratory and the other fixed on the moving system. The laboratory frame or, in general, any inertial frame used for observation can be termed as a *fixed frame*. A moving frame of reference, in general, can possess a translational or a rotational velocity with respect to a fixed frame of reference. Sometimes, a moving frame may possess both translational and rotational velocities relative to a fixed frame of reference. A frame of reference moving with a constant velocity relative to a fixed frame is called an *inertial* frame of reference. If, however, a frame of reference is accelerated relative to a fixed frame, it is called a *non-inertial* frame of reference. It is well known that rotational motion is always an accelerated motion. Hence, all frames of reference that are rotating relative to a fixed frame of reference are the non-inertial frames of reference.

The motion of a particle moving on the earth is usually described with reference to a frame fixed to the surface or to the centre of the earth. This frame is really speaking a non-inertial frame of reference because of the earth's rotational motion. In the analysis of the motion of a rigid body, rotating coordinate systems are found to be useful.

9.1 COORDINATE SYSTEMS WITH RELATIVE TRANSLATIONAL MOTIONS

Let us consider two coordinate systems $O'(x'y'z')$ fixed in space (fixed frame) and $O(xyz)$ fixed to a certain body that is moving with a translational velocity with respect to the first system or frame (Fig. 9.1).

Let $\mathbf{R}$ be the position vector of the point O with respect to the point O' at a certain instant t. Let P be any point whose position vectors with

respect to O and O' are $\mathbf{r}$ and $\mathbf{r}'$ respectively. Then, from Fig. 9.1, we have

$$\mathbf{r}' = \mathbf{R} + \mathbf{r} \tag{9.1}$$

Differentiating equation (9.1) with respect to time, we get

$$\dot{\mathbf{r}}' = \dot{\mathbf{R}} + \dot{\mathbf{r}} \tag{9.2}$$

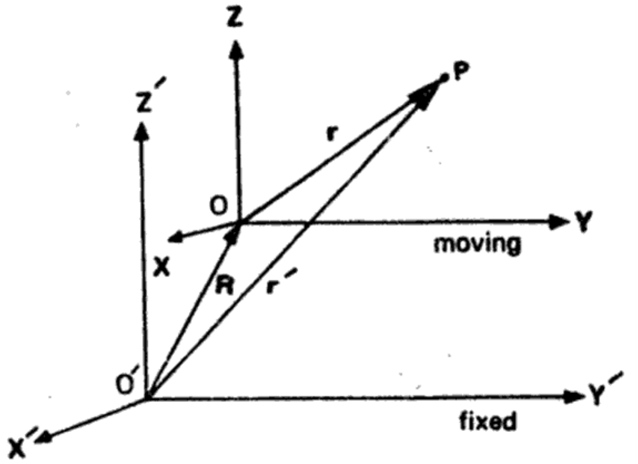


Fig. 9.1 Coordinate systems in relative uniform translatory motion

Thus, the velocity $\dot{\mathbf{r}}'$ of a particle at the point P measured in the fixed system is the vector addition of the velocity $\dot{\mathbf{R}}$ of the moving frame and the velocity $\dot{\mathbf{r}}$ of the particle at the point P in the moving frame. The corresponding accelerations are given by

$$\ddot{\mathbf{r}}' = \ddot{\mathbf{R}} + \ddot{\mathbf{r}} \tag{9.3}$$

The equation of motion of the particle at the point P in the fixed frame is

$$m\ddot{\mathbf{r}}' = \mathbf{F} \tag{9.4}$$

where $\mathbf{F}$ is the total external force acting on the particle P. The equation of motion of the particle at the point P in the moving frame of reference is

$$m\ddot{\mathbf{r}} = m\ddot{\mathbf{r}}' - m\ddot{\mathbf{R}}$$

$$= \mathbf{F} - m\ddot{\mathbf{R}}$$

or

$$m\ddot{\mathbf{r}} = \mathbf{F}_{\text{eff}} \tag{9.5}$$

Thus, if the moving frame of reference has an acceleration $\ddot{\mathbf{R}}$, the effective force acting on the particle at point P is smaller than the actual force by an amount $m\ddot{\mathbf{R}}$. This reduction arises as a result of the acceleration $\ddot{\mathbf{R}}$ of the moving frame. When $\ddot{\mathbf{R}} = 0$, the equations of motion are identical in the two systems. In other words, Newton's laws of motion are valid in the two systems moving with a uniform relative velocity. This is known as the *principle of Newtonian relativity* or *Galilean invariance*. The form of equations remains the *same* in the two systems. This is expressed by saying that the equations are *covariant* with respect to uniform translation of the coordinate systems. Such unaccelerated coordinate systems are called the *inertial frames of reference*.

According to the theory of relativity, however, the concepts of absolute rest or absolute motion are meaningless and hence an inertial frame which is at absolute rest or in absolute uniform motion cannot be differentiated.

Thus, all inertial frames are equivalent in the sense that the equations of motion are covariant in form. We still define, in classical mechanics, a frame of reference that approximates to an inertial frame at absolute rest by neglecting the effects that arise due to its motion. For example, a frame of reference fixed in the laboratory is used to study the rotational and precessional motion of the bodies even though this frame is moving along with the earth. The approximation will be valid if the effects due to the earth's motion are negligible as compared to those in the motion under study.

The terms 'coordinate system' and 'frames of reference' are generally used synonymously. Sometimes a slight difference is made in the meaning of these terms. In changing the coordinate systems, changes from one system to another that do not involve change of time are considered, while changes in the frames of reference include change in time. Thus, the frame of reference includes all the coordinate systems at rest with respect to any particular system. We shall, however, use both the terms synonymously.

If we are using a moving frame of reference, a term occurs in the equation of motion due to the acceleration of the frame. If we wish to get the same form for the equation of motion, we have to add a term $(-m\ddot{\mathbf{R}})$ to the right-hand side. This additional term does not represent any real force. It is called a *fictitious* or a *pseudo* or a *non-inertial force*. A real force has its existence due to certain kind of field of force or an interaction and it depends upon the position and motion of other bodies. The 'non-inertial force' has its existence only in a moving frame of reference, for example, the centrifugal force which is experienced by a person sitting in a car that turns suddenly to the left. The centrifugal force acts to the right.

9.2 ROTATING COORDINATE SYSTEMS

Let us suppose that the unprimed coordinate system $-O(x, y, z)$ is rotating with angular velocity $\boldsymbol{\omega}$ about some instantaneous axis passing through the origin O. The two systems, viz. the fixed and the rotating systems, have the common origin O. The fixed system is, as before, the primed system $O'(x', y', z')$ (Fig. 9.2).

The unit vectors $\mathbf{i}'$, $\mathbf{j}'$ and $\mathbf{k}'$ in the primed system are constant unit vectors whereas the unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ in the unprimed system are changing their directions along with the rotating axes.

The position vector of a particle at P can be written as

$$\left.\begin{aligned} \mathbf{r} &= \mathbf{i}'x' + \mathbf{j}'y' + \mathbf{k}'z' \\ \text{and also} \qquad \mathbf{r} &= \mathbf{i}x + \mathbf{j}y + \mathbf{k}z \end{aligned}\right\} \tag{9.6}$$

The transformation equations from unprimed system to primed system can be obtained by taking the dot product of $\mathbf{r}$ with $\mathbf{i}'$, $\mathbf{j}'$ and $\mathbf{k}'$. These

are

$$\left.\begin{aligned} x' &= (\mathbf{r}\cdot\mathbf{i}') = \mathbf{i}\cdot\mathbf{i}'x + \mathbf{j}\cdot\mathbf{i}'y + \mathbf{k}\cdot\mathbf{i}'z \\ y' &= (\mathbf{r}\cdot\mathbf{j}') = \mathbf{i}\cdot\mathbf{j}'x + \mathbf{j}\cdot\mathbf{j}'y + \mathbf{k}\cdot\mathbf{j}'z \\ z' &= (\mathbf{r}\cdot\mathbf{k}') = \mathbf{i}\cdot\mathbf{k}'x + \mathbf{j}\cdot\mathbf{k}'y + \mathbf{k}\cdot\mathbf{k}'z \end{aligned}\right\} \tag{9.7}$$

and

The dot products on the right-hand sides of the equation (9.7), viz. $\mathbf{i}\cdot\mathbf{i}'$, $\mathbf{j}\cdot\mathbf{i}'$, . . . , etc. are simply the cosines of the angles between the corresponding axes. The equations for inverse transformations can be similarly written down by taking the dot products of $\mathbf{r}$ with $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$.

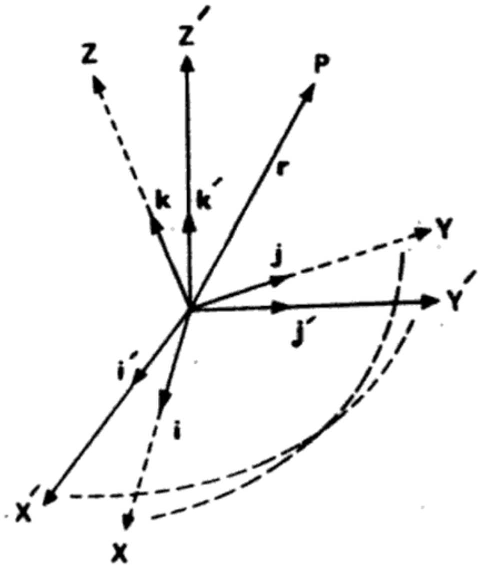


Fig. 9.2 Rotating coordinate axes

It is always possible to obtain equations similar to equations (9.6) and (9.7) for any vector $\mathbf{V}$ not necessarily drawn through the origin. Thus, a vector function $\mathbf{V} = \mathbf{V}(t)$ can be written as

$$\mathbf{V} = \mathbf{i}V_x + \mathbf{j}V_y + \mathbf{k}V_z = \mathbf{i}'V_x' + \mathbf{j}'V_y' + kV_z' \tag{9.8}$$

The time derivatives of $\mathbf{V}$, however, will be different in the two systems. In the primed, i.e. the fixed system, we have

$$\left(\frac{d\mathbf{V}}{dt}\right)_{\text{fix}} = \mathbf{i}'\dot{V}_x' + \mathbf{j}'\dot{V}_y' + \mathbf{k}'\dot{V}_z' \tag{9.9}$$

since the unit vectors are constant vectors.

In the unprimed, i.e., the rotating system, however, the unit vectors are changing in directions, hence their time derivatives will appear in the expression for $\frac{d\mathbf{V}}{dt}$. Thus

$$\left(\frac{d\mathbf{V}}{dt}\right)_{\text{fix}} = \mathbf{i}\dot{V}_x + \mathbf{j}\dot{V}_y + \mathbf{k}\dot{V}_z + \frac{d\mathbf{i}}{dt}V_x + \frac{d\mathbf{j}}{dt}V_y + \frac{d\mathbf{k}}{dt}V_z \tag{9.10}$$

The first three terms on the right-hand side of equation (9.10) are the time derivatives of the vector in the rotating system when the unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are treated as constant unit vectors.

Thus,

$$\left(\frac{d\mathbf{V}}{dt}\right)_{\text{rot}} = \mathbf{i}\dot{V}_x + \mathbf{j}\dot{V}_y + \mathbf{k}\dot{V}_z = \dot{\mathbf{V}}_r \tag{9.11}$$

represents the velocity in the rotating system.

The remaining three terms on the right-hand side of equation (9.10) arise as a result of the rotation of the system.

Now, the linear velocity of a particle having a position vector $\mathbf{r}$ and rotating with angular velocity $\boldsymbol{\omega}$ about the axis passing through the same origin is given by equation (1.77) as

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \boldsymbol{\omega} \times \mathbf{r} \tag{9.12}$$

In this expression, $\mathbf{r}$ is any position vector rotating in the body and the left-hand side is the time derivative of $\mathbf{r}$. This formula, therefore, can be applied to unit vectors as a special case.

Since, $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are the unit vectors in a system rotating with angular velocity $\boldsymbol{\omega}$, we have

$$\frac{d\mathbf{i}}{dt} = \boldsymbol{\omega} \times \mathbf{i},\ \frac{d\mathbf{j}}{dt} = \boldsymbol{\omega} \times \mathbf{j} \text{ and } \frac{d\mathbf{k}}{dt} = \boldsymbol{\omega} \times \mathbf{k} \tag{9.13}$$

Hence, equation (9.10) reduces to

$$\left(\frac{d\mathbf{V}}{dt}\right)_{\text{fix}} = \left(\frac{d\mathbf{V}}{dt}\right)_{\text{rot}} + \boldsymbol{\omega} \times \mathbf{V} \tag{9.14}$$

Equation (9.14) can be treated as an operator equation which gives the relations between the time derivatives in the fixed and the rotating coordinate systems. Thus

$$\left(\frac{d}{dt}\right)_{\text{fix}} = \left(\frac{d}{dt}\right)_{\text{rot}} + \boldsymbol{\omega} \times \tag{9.15}$$

The operator of equation (9.15) can be operated on any vector. If, in particular, we operate it on $\boldsymbol{\omega}$, we get

$$\begin{aligned}\left(\frac{d\boldsymbol{\omega}}{dt}\right)_{\text{fix}} &= \left(\frac{d\boldsymbol{\omega}}{dt}\right)_{\text{rot}} + \boldsymbol{\omega} \times \boldsymbol{\omega} \\ &= \left(\frac{d\boldsymbol{\omega}}{dt}\right)_{\text{rot}} \\ &= \dot{\boldsymbol{\omega}}\end{aligned} \tag{9.16}$$

since $\boldsymbol{\omega} \times \boldsymbol{\omega} = 0$.

Equation (9.16) shows that the angular acceleration $\dot{\boldsymbol{\omega}}$ is the same in the fixed and the rotating systems.

The second derivative of $\mathbf{V}$ can be found out in a similar manner. To simplify the notation, let us denote

$$\left(\frac{d}{dt}\right)_{\text{fix}} = \frac{d'}{dt} \text{ and } \left(\frac{d}{dt}\right)_{\text{rot}} = \frac{d}{dt} \tag{9.17}$$

situated at point P is directed towards the axis of rotation and is perpendicular to it. It has a magnitude

$$|\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})| = \omega^2 r \sin \theta = \frac{v^2}{r \sin \theta} \tag{9.22}$$

where $v = \omega r \sin \theta$ is the speed of the particle when it rotates in a circle of radius $r \sin \theta$ (Fig. 9.3). The Coriolis acceleration is present only when the particle has a velocity $\dot{\mathbf{r}}_r$ in the rotating frame.

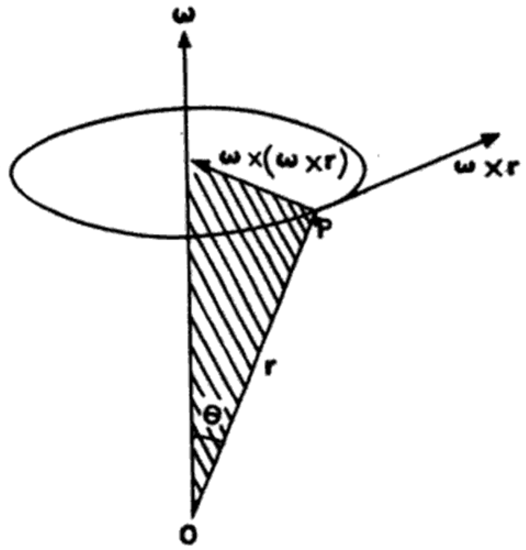


Fig. 9.3 Centripetal acceleration

9.3 THE CORIOLIS FORCE

We have earlier seen that Newton's second law of motion, viz. $\mathbf{F} = m\ddot{\mathbf{r}}$ is valid only in the inertial frames of reference. Hence,

$$\mathbf{F} = m\left(\frac{d^2\mathbf{r}}{dt^2}\right)_{\text{fix}} \tag{9.23}$$

In this equation, the differentiation must be carried out with respect to the fixed system. We wish to write the equation of motion of a particle in the rotating system such that it has the same form. Let the angular velocity of the rotating system be constant. Then, $\dot{\boldsymbol{\omega}} = 0$. Let, further, the origins of the fixed and the moving systems coincide. Then, $\mathbf{R} = 0$ and we have $\mathbf{r}' = \mathbf{r}$. Equation (9.21) then gives

$$\begin{aligned} m\left(\frac{d^2\mathbf{r}}{dt^2}\right)_{\text{rot}} &= m\left(\frac{d^2\mathbf{r}}{dt^2}\right)_{\text{fix}} - 2m\boldsymbol{\omega} \times \left(\frac{d\mathbf{r}}{dt}\right)_{\text{rot}} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) \\ &= \mathbf{F}_{\text{eff}} \end{aligned} \tag{9.24}$$

Thus, the forces acting on the particle in the rotating frame are (i) the real force $\mathbf{F} = m\left(\dfrac{d^2\mathbf{r}}{dt^2}\right)_{\text{fix}}$, (ii) the centrifugal force $-m\boldsymbol{\omega}(\boldsymbol{\omega} \times \mathbf{r})$ arising as a result of the rotation of coordinate axes, and (iii) the Coriolis force $-2m\boldsymbol{\omega} \times \left(\dfrac{d\mathbf{r}}{dt}\right)_{\text{rot}}$ arising as a result of the motion of the particle in the rota-

ting system. It should be noted that the centrifugal force and the Coriolis force are arising because we have used a non-inertial frame. These are added to the force term $\mathbf{F} = m\left(\frac{d^2\mathbf{r}}{dt^2}\right)_{\text{rot}}$ so as to give $\mathbf{F}_{\text{eff}}$. Hence

$$\mathbf{F} + \text{non-inertial forces} = \mathbf{F}_{\text{eff}} \tag{9.25}$$

This is what we shall have to do if we wish to write the equation of motion of a particle in the rotating frame so that the equation of motion resembles Newton's law, viz. $\mathbf{F} = m\ddot{\mathbf{r}}$. Thus, for a satellite moving around the earth in the fixed system, the only real force acting on it is the gravitational force that produces the centripetal acceleration. An observer in the satellite, however, will experience only the effective force

$$\mathbf{F}_{\text{eff}} = \mathbf{F} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})$$

As the observer is at rest with respect to the revolving satellite, we have to postulate, as we observe from the earth, a fictitious force, viz. the centrifugal force in order to balance the gravitational force. Just as an observer in a freely falling lift will not experience any force and will be in a weightless condition, an observer in a revolving satellite will also be in a weightless condition.

The noninertial forces are often helpful in dealing with the problems concerned with rotational motion. The use of rotating coordinate systems may appear complicated; but in these systems the problems are simplified considerably and brought to the form of Newton's equation, viz. $\mathbf{F} = m\ddot{\mathbf{r}}$.

9.4 MOTION ON THE EARTH

Let us fix two coordinate systems—one at the centre of the earth but fixed in space (neglecting translational motion of the earth) and the other at some point in the body of the earth but rotating along with the earth with an angular velocity $\boldsymbol{\omega}$. A particle of mass m situated on the surface of the earth will be acted upon by the gravitational force $m\mathbf{g}$. Some other real force $\mathbf{F}$—such as an electrostatic or a magnetic force—may also act on it. Then, the equation of motion of the particle in the fixed system is

$$m\left(\frac{d^2\mathbf{r}}{dt^2}\right)_{\text{fix}} = \mathbf{F} + m\mathbf{g} \tag{9.26}$$

The equation of motion of the same particle in the rotating system is

$$m\left(\frac{d^2\mathbf{r}}{dt}\right)_{\text{rot}} = \mathbf{F} + m[\mathbf{g} - \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})] - 2\boldsymbol{\omega} \times \left(\frac{d\mathbf{r}}{dt}\right)_{\text{rot}} \tag{9.27}$$

The second term on the right-hand side of equation (9.27) represents the effective gravitational acceleration

$$\mathbf{g}_e = \mathbf{g} - \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) \tag{9.28}$$

The gravitational acceleration measured at any point will be this effective acceleration and it will be less than the acceleration due to the earth if it were not rotating. The term $-\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})$ is called the centrifugal acceleration (Fig. 9.4). It always points radially outwards. The centrifugal acceleration will be found to be zero at the poles, since at the

component of $\boldsymbol{\omega}$ along the vertical direction is $\omega_z\hat{\mathbf{e}}_z = \hat{\mathbf{e}}_z\omega \sin \theta$ and is directed upward in the northern hemisphere and downward in the

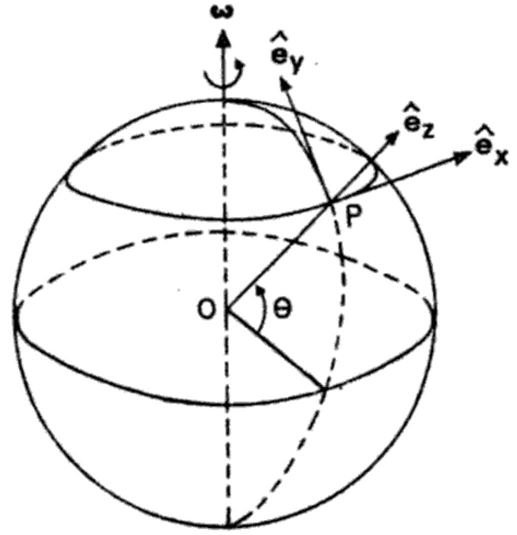


Fig. 9.5 Particle P has latitude θ

southern hemisphere. Hence, the path of the particle will be deflected towards the right in the northern hemisphere and towards the left in the southern hemisphere due to the Coriolis acceleration. The maximum magnitude of the Coriolis acceleration is at the north pole (Fig. 9.6), or the south pole and is

$$2\omega v_r = 1.5 \times 10^{-4} v_r$$

where v_r is the velocity in the horizontal plane.

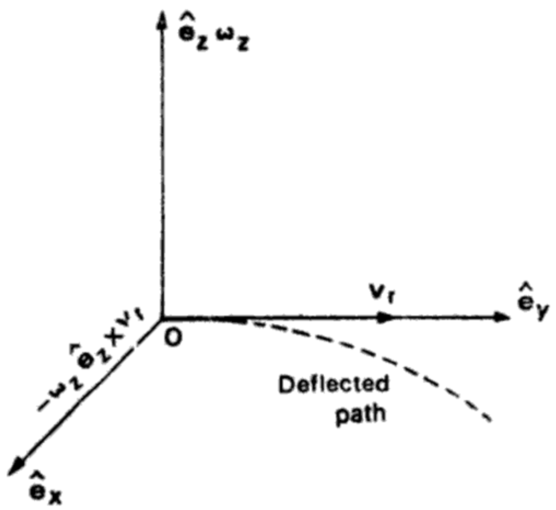


Fig. 9.6 Deflection of the path of a particle due to Coriolis force

For velocity of 1 km/s or 3600 km/hr, the magnitude of the Coriolis acceleration is 0.15 m/s^2, which is about 0.015 g.

The component of $\boldsymbol{\omega}$ along a vertical direction in a local coordinate system at the equator will be zero. Hence, the Coriolis force acting on a horizontally moving particle will be zero.

Although the magnitude of the Coriolis acceleration is small, it plays an important role in many phenomena on the earth.

It is important to take into consideration the effects of the Coriolis acceleration in the flight of missiles, the velocity and the time of flight of which are considerably large.

If the velocity of the particle is directed towards the north in the northern hemisphere, then the deflection due to the Coriolis force is towards the east. The deflection is sufficiently small and the angle of deflection from Figs. 9.5 and 9.6 is

$$\alpha = \frac{\text{distance travelled in time } t \text{ in the deflected direction}}{\text{distance travelled in time } t \text{ in the direction of projection}}$$

$$= \frac{\frac{1}{2}(2\omega_z v_r)t^2}{v_r t}$$

$$= \omega_z t = \omega \sin\theta\, t \tag{9.29}$$

Thus, on the north pole $\theta = 90°$ and the deflection α due to the Coriolis acceleration is simply the angle of rotation of the earth in a time t. The value of α is maximum at the north pole and zero at the equator. If $t = 100$ s, the maximum deflection of a missile projected from a point on the equator is $\alpha = 7\times10^{-4}$ radian $= 0.04°$. Thus, the deflection is found to be quite small, but it assumes considerable importance in guided missiles.

Another terrestrial phenomenon in which the Coriolis force plays an important role is the formation of cyclones. Whenever a low pressure region is formed, a mass of air will rush to this region from all directions. But, due to the Coriolis force, it will be deflected to the right in the Northern Hemisphere (Fig. 9.7). Thus, in a cyclone the wind whirls

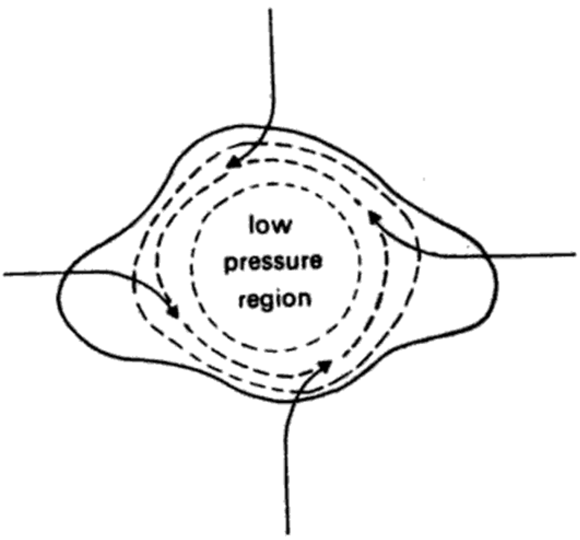


Fig. 9.7 The wind whirls in the anticlockwise sense in the northern hemisphere.

in the anticlockwise sense in the northern hemisphere. The sense of rotation of wind will be clockwise in the southern hemisphere.

9.5 EFFECT OF CORIOLIS FORCE ON A FREELY FALLING PARTICLE

Consider a particle which is falling freely towards the earth. Let the height h through which the particle falls be so small that the variation in g can be neglected. The acceleration of the particle is then given by

$$\mathbf{a} = \mathbf{g} - 2\boldsymbol{\omega} \times \mathbf{v} \tag{9.30}$$

where $\mathbf{a}$ and $\mathbf{v}$ are measured with respect to the earth, i.e. in the rotating frame. From Fig. 9.5, in the northern hemisphere, we have

$$\omega_x = 0,\ \omega_y = \omega \cos\theta \text{ and } \omega_z = \omega \sin\theta$$

The deflection produced by the Coriolis force is quite small and hence for a particle moving along $-\hat{\mathbf{e}}_z$ direction, the force will have negligibly small components along the directions $\hat{\mathbf{e}}_x$ and $\hat{\mathbf{e}}_y$. Thus

$$\dot{x} \simeq 0, \quad \dot{y} \simeq 0 \quad \text{and} \quad \dot{z} = -gt$$

Hence, $\boldsymbol{\omega} \times \mathbf{v} = -\omega gt \cos\theta\, \hat{\mathbf{e}}_x$ and lies along the x-direction. As $\mathbf{g}$ is directed along the $-\hat{\mathbf{e}}_z$ direction, the components of the acceleration of the particle are

$$a_x = \ddot{x} = 2\omega gt \cos\theta \tag{9.31}$$

$$a_y = \ddot{y} = 0 \tag{9.32}$$

and

$$a_z = \ddot{z} = -g \tag{9.33}$$

Thus, from equation (9.31) the acceleration along the $\hat{\mathbf{e}}_x$ direction is seen to be due to the Coriolis force. Integrating equations (9.31) and (9.33), we get the solutions

$$x = \tfrac{1}{3}\omega gt^3 \cos\theta \tag{9.34}$$

and

$$z = z_0 - \tfrac{1}{2}gt^2 \tag{9.35}$$

where the initial position is $z(0) = Z_0$ and $x(0) = 0$ and the initial velocity is $\dot{x}(0) = \dot{z}(0) = 0$.

Now, from equation (9.35), the time of fall from height h is

$$t = \sqrt{\frac{2h}{g}} \tag{9.36}$$

Hence the deflection of a particle towards the east when it is dropped from rest is given by

$$\tfrac{1}{3}\omega \cos\theta \left(\frac{8h^3}{g}\right)^{1/2} \tag{9.37}$$

If a particle is dropped from a height of 100 m from rest at latitude $\theta = 45°$, it will be deflected by about 1.55×10^{-2} m. In the treatment given above, the effects of wind, viscosity etc. are neglected.

QUESTIONS

1. Is the earth really an inertial frame of reference? Explain.
2. What is a non-inertial frame? Give illustrations.

5. Water is standing in a tub. When the drain hole is opened will the flow of water develop a vortex ? What will be its sense of rotation in the northern hemisphere ?
6. A pail and the water in it are rotating about a vertical axis. Find the equation for the water surface, neglecting surface tension.
7. A balloon at altitude 10 km and latitude 60° drops a bag of ballast. Neglecting air resistance, what will be the horizontal displacement of the bag during its fall ?
8. Two identical cylinders are attached to a horizontal turntable. The centre of each is at distance R from the axis of rotation. One cylinder contains air and a pingpong ball hanging from a thread. The other contains water with a pingpong ball attached to the bottom with a string. If the angular velocity is ω what are the directions of the strings ?
9. Show that the plane of oscillation of a pendulum at latitude θ rotates through $2\pi \sin \theta$ every day. (Foucault pendulum). Give the physical explanation of this for a pendulum on a pole of the earth.
10. A canon of range 10 km is at latitude 60°, pointing due south. What is the east-west displacement of the canon-ball at the end of its flight ?
11. A bead slides along a smooth curved wire given by $z = c(x^2 + y^2)$. It rotates about the axis with angular velocity ω. The z-axis is vertical. Including gravity, what is the equation of motion ? At what height can the bead remain stationary ?
12. Consider a particle sliding inside a smooth paraboloid $z = C(x^2+y^2)$, under the action of gravity. (a) Find the equation for z. (b) Find the value of z corresponding to a horizontal circular motion with $\varphi = \omega$. (c) Compare with the answer of problem 11. Why do the two agree, although the forces involved are, in general, different ? Show that p_φ is constant in this problem but not in problem 11.
13. In the above problem, once the speed v is given and the path is required to be a circle, $z = 0$ and the value of z is determined by the given formula. If these conditions are not satisfied, the actual path oscillates about the circular one. Show that, when the disturbance is small, the oscillations are simple harmonic (use rotating frame). Therefore, since the average path is still the same circle, the circular motion is said to be stable. The study of stability of circular orbit is very important in designing accelerators.
14. A projectile is fired horizontally along the earth's surface. Show that to a first approximation the angular deviation from the direction in which the projectile is fired resulting from the Coriolis force varies linearly with time at a rate $\omega \cos \theta$, where ω is the angular frequency of the earth's rotation, θ is the colatitude, the direction of deviation being to the right in the northern hemisphere.

15. A body is dropped from a height h above the surface of the earth. (a) Calculate the Coriolis acceleration as a function of time t, assuming that the distance y of the body from the surface of the earth as a function of time is given by $y = h - \frac{1}{2}gt^2$. (b) Compute the net displacement d of the point of impact due to the earth's rotation. (Assume the initial velocity of the body to have been zero with respect to the centre of the earth).
16. Calculate the centrifugal acceleration, due to the earth's rotation, of a particle at the equator and on the surface of the earth. Compare this with g. Also find the centrifugal acceleration due to the motion of the earth around the sun. Verify that this acceleration may be neglected in comparison with the acceleration due to axial rotation.
17. Show that a particle, projected vertically upward from a point on the earth's surface at a northern latitude λ strikes the ground at a point

$$\tfrac{4}{3}\omega \cos \lambda \sqrt{8h^3/g}$$

to the west. Neglect air resistance and consider only small vertical heights.
18. If a projectile is fired due east from a point on the surface of the earth at a northern latitude λ with velocity V_0 and at angle α with the horizontal, show that the lateral deflection when the projectile strikes the earth is

$$d = \frac{4V_0^3}{g^2} \omega \sin \lambda \sin^2 \alpha \cos \alpha$$

where ω is the rotation frequency of the earth.
19. In problem 18, if R represents the range of the projectile if the earth did not rotate, show that the change in the range due to the rotation of the earth is

$$\Delta R = \sqrt{\frac{2R^3}{g}}\, \omega \cos \lambda\, [\cot^{1/2} \alpha - \tfrac{1}{3} \tan^{3/2} \alpha]$$

20. Show that the angular deviation of a plumb-line from the true vertical at a point on the earth's surface at a latitude λ is

$$\frac{r_0\omega^2 \sin \lambda \cos \lambda}{g - r_0\omega^2 \cos^2 \lambda}$$

where r_0 is the radius of the earth.
21. An object is thrown downward with initial speed v_0; prove that after a time t, it is deflected east of the vertical by an amount

$$\omega v_0 \sin \lambda t + \tfrac{1}{3}\omega g t^3 \sin \lambda$$

22. A particle is thrown vertically upward at a colatitude λ with speed v_0. Prove that when it returns it will be at a distance

$$\frac{4\omega v_0^3 \sin \lambda}{3g^2}$$

eastward from the starting point.

23. An object at the equator is thrown vertically upward with a speed of 60 mph. How far from its original position will it land?
24. A vertical rod MN is rotating with constant angular velocity ω. An inextensible light string of length l has one of its ends attached to a point O on the rod. The other end P carries a mass M attached to it. Find the tension in the string and the angle made by OP with the vertical in the equilibrium position.
25. How long would it take for the plane of oscillation of a Foucault pendulum to make one complete revolution if the pendulum is located (a) at the north pole, and (b) at a place of colatitude 45°?

10
Motion of a Rigid Body

A body is said to be a rigid body if the distance between any two of its constituent particles is a constant. However, no material body in nature is perfectly rigid in the strict sense of the word. Moreover, the particles constituting a body are never at rest. All the bodies are made up of atoms and molecules which are in turn composed of sub-atomic particles. These particles are always in a state of incessant motion. This motion of the microscopic particles can always be neglected while studying the motion of the rigid bodies. There are other types of displacements such as the elastic deformations of the body in which the size and the shape of the body change. These microscopic changes are quite small in some types of materials and hence these can also be neglected and the given body can be treated as a rigid body.

A rigid body can have two types of motion—a *translational* motion and a *rotational* motion. A rigid body in motion can, therefore, be completely specified if its position and orientation are given. For this purpose, we must specify the coordinates of any three points not in the same straight line in the body. Because if the body is fixed at one point, it can rotate about any axis passing through that point. If one more point of the body is fixed, the body can rotate about the axis passing through the two fixed points. Then, the coordinates of the third point not on the axis, alone will be able to locate the rigid body completely in space.

Any point in space is specified in terms of its three coordinates in some convenient coordinate system. Hence, for the location of a rigid body, we would need nine coordinates. But these nine coordinates are not independent of each other and there exist three equations of the constraints, viz.

$$(\mathbf{r}_i - \mathbf{r}_j)^2 = \text{a const} \tag{10.1}$$

where $\mathbf{r}_i$ or $\mathbf{r}_j$ (i or $j = 1, 2, 3, \ldots$) represent the position vectors of the

three points in the rigid body. Hence, only six coordinates are independent of each other. Thus, a rigid body is said to have six degrees of freedom.

The six degrees of freedom of a rigid body can be interpreted in yet another way. Three independent coordinates will be required to locate a point on the instantaneous axis of rotation in the rigid body which is undergoing a translational motion relative to some fixed or inertial frame of reference. Two more coordinates viz. the direction cosines will be required to locate the axis of rotation passing through the point already located. Lastly, the orientation of the body can be specified in terms of an angle. Thus, out of the six degrees of freedom, three degrees of freedom represent the free translational motion of the body whereas the remaining three degrees of freedom represent its rotational motion. It should be noted that the orientation of a rotating body can be expressed in terms of three angles. One possible choice of these angles is called Euler's angles (article 10.6).

In the discussion of the motion of a rigid body, we shall use two coordinate systems—one of which is fixed in space, i.e. an inertial frame, and the other fixed in the body. The latter system is referred to as the *body-coordinate system* and is a non-inertial frame of reference. This is because this system undergoes a translational and a rotational motion along with the rigid body.

10.1 EULER'S THEOREM

Euler's theorem is one of the basic theorems used in the description of the motion of a rigid body. It states that *any general displacement of a rigid body, one point of which is fixed, is a rotation about some axis passing through the fixed point.*

As one point of the rigid body is fixed, the body does not have any translational motion. Hence, according to the above theorem we can always find out a single rotation about some axis for any arbitrary rotation from its original to final orientation.

Let us take the *body-system of coordinates* such that the origin of the system coincides with the fixed point. In the rotational motion, the position vector of any particle in the body does not change in its magnitude. The theorem would be proved if we can find a straight line—i.e. the axis of rotation—such that the distance of the particles of the body from this straight line remains constant during the rotation.

Let A and B be the initial positions of two particles in the rigid body. These particles occupy positions A' and B' after some arbitrary rotational displacement. Let O be the fixed point (Fig. 10.1). Thus, the body was in the initial configuration OAB and its final configuration is $OA'B'$. Let us draw two planes P_1 and P_2 perpendicular to the planes of the triangles OAA' and OBB' respectively. Let the planes divide the angles of the triangles at O and intersect each other along the straight line OC. Now,

every point on the plane P_1 is equidistant from points A and A'. Similarly, points B and B' are equidistant from plane P_2. The straight line OC

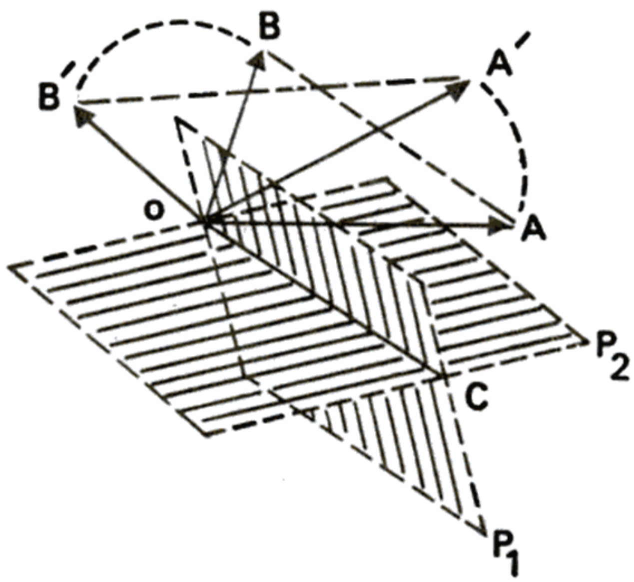


Fig. 10.1 Euler's theorem

of the interaction of planes P_1 and P_2 is such that the distances of particles A and B from the straight line OC before and after rotation are equal. Thus, the straight line OC is the axis of rotation. Even when the body is brought from the initial configuration OAB to its final configuration $OA'B'$, OC remains unchanged and the displacement is equivalent to a rotation about OC.

The above considerations help us to understand another important theorem—*Chasles' theorem.* It states that *the most general displacement of the rigid body is a translation plus a rotation about some axis.* Chasles' theorem leads us to the idea that it is possible to separate the discussion of the motion of the rigid bodies into two parts—one dealing with the translational motion and the other dealing with the rotational motion. This fact is also brought out in the division of the six degrees of motion mentioned above. Such a separation is always possible and was obtained in Chapter 3. If we choose the fixed point to be at the centre of mass of the body, the total angular momentum is equal to the sum of the angular momentum of the centre of mass and the angular momenta of the constituent particles about the centre of mass. A similar relation was obtained in the case of the kinetic energy.

Thus, the separation of the discussion of the motion of a rigid body into the translational and the rotational motion helps us to study the two aspects of the motion separately. Since, we have studied some aspects of the translational motion earlier, we shall, in this chapter, concentrate on the rotational motion only.

10.2 ANGULAR MOMENTUM AND KINETIC ENERGY

Consider a rigid body composed of n particles having masses m_a $(a = 1, 2, \ldots, n)$ and rotating with instantaneous angular velocity $\boldsymbol{\omega}$. Let one of the points in the body be fixed. Hence, translational motion is absent. We shall now find out the expressions for the angular momentum and the kinetic energy due to the rotation of the body.

The linear velocity $\mathbf{v}_a$ of the particle of mass m_a and position vector $\mathbf{r}_a$ with respect to the fixed point is given by

$$\mathbf{v}_a = \boldsymbol{\omega} \times \mathbf{r}_a \tag{10.2}$$

The total angular momentum $\mathbf{L}$ is the sum of angular momenta $\boldsymbol{l}_a$ of the individual particles and is given by

$$\begin{aligned} \mathbf{L} = \sum_{a=1}^{n} \boldsymbol{l}_a &= \sum_a \mathbf{r}_a \times m_a \mathbf{v}_a \\ &= \sum_a m_a \mathbf{r}_a \times (\boldsymbol{\omega} \times \mathbf{r}_a) \\ &= \sum_a m_a r_a^2 \boldsymbol{\omega} - \sum_a m_a (\mathbf{r}_a \cdot \boldsymbol{\omega}) \mathbf{r}_a \end{aligned} \tag{10.3}$$

The summation is carried out over all the particles of the rigid body. Equation (10.3) can be written in component form. Thus, the x-component of the angular momentum is given by

$$L_x = \sum_a m_a (r_a^2 - x_a^2) \omega_x - \sum_a m_a x_a y_a \omega_y - \sum_a m_a x_a z_a \omega_z \tag{10.4}$$

We now introduce the following symbols for the coefficients of ω_x, ω_y and ω_z.

$$I_{xx} = \sum_a m_a (r_a^2 - x_a^2) = \sum_a m_a (y_a^2 + z_a^2) \tag{10.5a}$$

$$I_{yy} = \sum_a m_a (z_a^2 + x_a^2) \tag{10.5b}$$

$$I_{zz} = \sum_a m_a (x_a^2 + y_a^2) \tag{10.5c}$$

$$I_{xy} = -\sum_a m_a x_a y_a = I_{yx} \tag{10.6a}$$

$$I_{yz} = -\sum_a m_a y_a z_a = I_{zy} \tag{10.6b}$$

and

$$I_{zx} = -\sum_a m_a z_a x_a = I_{xz} \tag{10.6c}$$

With these substitutions, equation (10.4) and the other components of the angular momenta can be written as

$$\left.\begin{aligned} L_x &= I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z \\ L_y &= I_{yx}\omega_x + I_{yy}\omega_y + I_{yz}\omega_z \\ \text{and} \quad L_z &= I_{zx}\omega_x + I_{zy}\omega_y + I_{zz}\omega_z \end{aligned}\right\} \tag{10.7}$$

The quantities I_{xx}, I_{yy} and I_{zz} are called the moments of inertia about x, y and z axes, respectively, and the quantities I_{xy}, I_{yz} and I_{zx} are called the products of inertia.

Equations (10.7) can be written in a compact form if we denote the x, y and z axes by 1, 2 and 3 respectively.

Thus, we have

$$L_i = \sum_j I_{ij}\omega_j,\ i = 1, 2, 3 \tag{10.8}$$

The coefficients I_{ij} can always be calculated if the distribution of particles about the axes is known.

The kinetic energy of the rigid body can be calculated on similar lines. We have

$$T = \sum_a \tfrac{1}{2} m_a v_a^2$$

or

$$\begin{aligned} 2T &= \sum_a m_a |\mathbf{v}_a|^2 \\ &= \sum_a m_a(\boldsymbol{\omega} \times \mathbf{r}_a)\cdot(\boldsymbol{\omega} \times \mathbf{r}_a) \\ &= \sum_a m_a \boldsymbol{\omega}\cdot[\mathbf{r}_a \times (\boldsymbol{\omega} \times \mathbf{r}_a)] \\ &= \boldsymbol{\omega}\cdot\sum_a m_a \mathbf{r}_a \times (\boldsymbol{\omega} \times \mathbf{r}_a) \\ &= \boldsymbol{\omega}\cdot\sum_a m_a \mathbf{r}_a \times \mathbf{v}_a \\ &= \boldsymbol{\omega}\cdot\mathbf{L} \end{aligned}$$

Thus,

$$2T = \boldsymbol{\omega}\cdot\mathbf{L} \tag{10.9}$$

In obtaining equation (10.9), we have used the property of interchange of the dot and the cross in the scalar triple product.

Thus, the kinetic energy T is given by

$$\begin{aligned} T &= \tfrac{1}{2}\boldsymbol{\omega}\cdot\mathbf{L} \\ &= \tfrac{1}{2}\sum_i \omega_i L_i \\ &= \tfrac{1}{2}\sum_{ij} I_{ij}\omega_i\omega_j \end{aligned} \tag{10.10a}$$

in which we have used $L_i = \sum_j I_{ij}\omega_j$ from equation (10.8).

On expansion, equation (10.10) becomes

$$T = \tfrac{1}{2}[I_{xx}\omega_x^2 + I_{yy}\omega_y^2 + I_{zz}\omega_z^2 + 2I_{xy}\omega_x\omega_y + 2I_{yz}\omega_y\omega_z + 2I_{zx}\omega_z\omega_x] \tag{10.10b}$$

If the body is rotating about the z-axis with the angular velocity $\boldsymbol{\omega}$, we have

$$\omega_z = \omega \text{ and } \omega_x = \omega_y = 0$$

Then, equation (10.10) becomes

$$T = \tfrac{1}{2}I_{zz}\omega_z^2 = \tfrac{1}{2}I\omega^2 \tag{10.11}$$

where I is the moment of inertia of the body about the z-axis. In this case, the components of the angular momenta are

$$L_x = I_{xz}\omega_z,\ L_y = I_{yz}\omega_z \text{ and } L_z = I_{zz}\omega_z \tag{10.12}$$

This shows that the directions of the angular velocity and the angular momentum are, in general, different.

Let us consider a simple system of two particles having masses m_1 and m_2 and connected by a rigid rod of negligible mass. The system is rotating with angular velocity $\boldsymbol{\omega}$ as shown in Fig. 10.2.

The angular momentum of the system is given by

$$\mathbf{L} = m_1\mathbf{r}_1 \times \mathbf{v}_1 + m_2\mathbf{r}_2 \times \mathbf{v}_2 \tag{10.13}$$

The angular momentum vector **L** must be perpendicular to the connecting rod and is clearly not parallel to the angular velocity vector $\boldsymbol{\omega}$. As the particles rotate, the vector **L** which lies in the shaded plane (Fig. 10.2) but is perpendicular to the rod, also rotates and traces a cone with vertex at the point O.

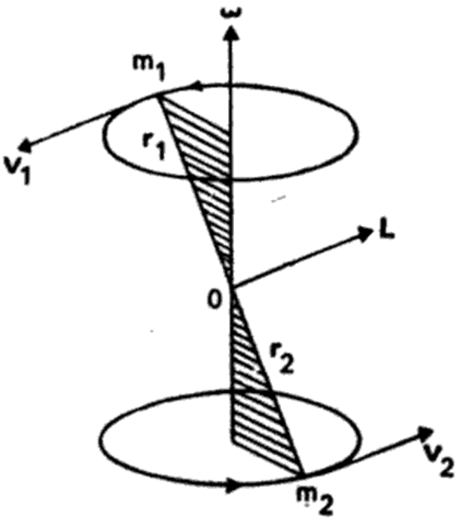


Fig. 10.2 The directions of the angular velocity vector and the angular momentum vector are different

Thus, $\dot{\mathbf{L}} \neq 0$ and according to the equation of motion

$$\dot{\mathbf{L}} = \mathbf{N} \tag{10.14}$$

Hence, a torque **N** must be applied to maintain the rotation of the system about the given axis.

10.3 THE INERTIA TENSOR

The nine terms I_{ij} mentioned earlier can be treated as the components of the moment of inertia of the rigid body. Each component is a scalar and has the dimension of $[M^1L^2]$. Since there are nine (i.e. 3^2) components, the physical quantity, moment of inertia, is a tensor of rank two. The moment of inertia tensor or the inertia tensor is denoted by $\overleftrightarrow{I}$. It can be written in a matrix notation, viz.

$$\overleftrightarrow{I} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \tag{10.15}$$

The moment of inertia tensor is symmetric, i.e.

$$I_{ij} = I_{ji}$$

and it has only six independent components.

If, in a body, the matter is continuously distributed and the densit

body for which $I_1 = I_2$ and $I_3 = 0$ is called a *rotor*, an example of which is a diatomic molecule.

10.4 EULER'S EQUATIONS OF MOTION

Consider a rigid body, one point of which is fixed. Let **N** be the torque acting on it. Then, the equation of the rotational motion of the rigid body in a fixed or inertial frame of reference is given by

$$\left(\frac{d\mathbf{L}}{dt}\right)_{\text{fix}} = \mathbf{N} \tag{10.23}$$

But, the time derivative in a fixed frame of reference is related to the time derivative in a rotating frame of reference by the following operator relation

$$\left(\frac{d}{dt}\right)_{\text{fix}} = \left(\frac{d}{dt}\right)_{\text{rot}} + \boldsymbol{\omega}\times$$

that operates on vectors. In the present case, the rigid body is rotating with angular velocity $\boldsymbol{\omega}$. We shall fix a frame of reference in it and call it a *body frame of reference*. Hence, equation (10.23) can be written as

$$\mathbf{N} = \left(\frac{d\mathbf{L}}{dt}\right)_{\text{fix}} = \left(\frac{d\mathbf{L}}{dt}\right)_{\text{rot}} + \boldsymbol{\omega}\times\mathbf{L} \tag{10.24}$$

The components of moment of inertia are, however, constant with respect to the body frame of reference. Hence, we have

$$\left(\frac{d\mathbf{L}}{dt}\right)_{\text{rot}} = \left(\frac{d\mathbf{L}}{dt}\right)_{\text{body}} = \left[\frac{d}{dt}(\overleftrightarrow{\mathbf{I}}\cdot\boldsymbol{\omega})\right]_{\text{body}} = \overleftrightarrow{\mathbf{I}}\cdot\dot{\boldsymbol{\omega}} \tag{10.25}$$

But, the derivative of $\boldsymbol{\omega}$ with respect to time is the same in the fixed and body frames of reference. Hence, using equation (10.25) and dropping the suffix, equation (10.24) becomes

$$\mathbf{N} = \overleftrightarrow{\mathbf{I}}\cdot\dot{\boldsymbol{\omega}} + \boldsymbol{\omega}\times\mathbf{L} \tag{10.26}$$

The expression for angular momentum will be greatly simplified if we orient the axes of the body frame of reference such that they coincide with the principal axes of the body. All the products of inertia will then vanish and equation (10.26) for the rigid body in the component form becomes

$$\left.\begin{aligned} N_1 &= I_1\dot{\omega}_1 + (I_3 - I_2)\omega_2\omega_3 \\ N_2 &= I_2\dot{\omega}_2 + (I_1 - I_3)\omega_3\omega_1 \\ \text{and}\qquad N_3 &= I_3\dot{\omega}_3 + (I_2 - I_1)\omega_1\omega_2 \end{aligned}\right\} \tag{10.27}$$

Equations (10.27) are known as the Euler's equations of the motion of a rigid body. The rigid body under consideration is constrained such that one point of it is fixed. This point itself is chosen as the origin. If the rigid body is not constrained, i.e., if it is moving freely, then the centre of mass of the body is chosen as the origin of the body frame of reference. It should be noted that the components of the torque and of the angular velocity in Euler's equations are taken along the axes of the body frame of reference which coincide with the principal axes at the origin.

The solutions of Euler's equations will enable us to understand how angular velocity of a rigid body changes with time with respect to the principal axes under the action of the known torque. Thus, we are in a position to know how vectors $\boldsymbol{\omega}$ and $\mathbf{L}$ move relative to the body axes.

If the torque acting on the rigid body is zero, then according to the law of conservation of angular momentum, the angular momentum of the body is a constant of the motion. Hence, the first term of the right-hand side of equation (10.24) is zero and we are left with

$$\boldsymbol{\omega} \times \mathbf{L} = 0$$

or

$$\boldsymbol{\omega} \times (\overleftrightarrow{\mathbf{I}} \cdot \boldsymbol{\omega}) = 0 \tag{10.28}$$

Equation (10.28) will be true only when the angular momentum vector $\mathbf{L}$ is parallel to the angular velocity vector $\boldsymbol{\omega}$, i.e., when

$$\mathbf{L} = I\boldsymbol{\omega} \tag{10.29}$$

where I is the scalar magnitude of the moment of inertia taken about one of the principal axes. Thus, in a torque-free motion of a rigid body, $\overleftrightarrow{\mathbf{I}} \cdot \boldsymbol{\omega}$ is parallel to $\boldsymbol{\omega}$, i.e., the angular velocity vector $\boldsymbol{\omega}$ is directed along the principal axis of the body. Let us now take the dot product of equation (10.26) with $\boldsymbol{\omega}$ and use equation (10.28) simultaneously. Then, we get

$$\begin{aligned}
\boldsymbol{\omega} \cdot \mathbf{N} &= \boldsymbol{\omega} \cdot \overleftrightarrow{\mathbf{I}} \cdot \frac{d\boldsymbol{\omega}}{dt}, \quad \text{since} \quad \boldsymbol{\omega} \cdot \boldsymbol{\omega} \times \mathbf{L} = 0 \\
&= \frac{d\boldsymbol{\omega}}{dt} \cdot \overleftrightarrow{\mathbf{I}} \cdot \boldsymbol{\omega} \\
&= \frac{1}{2} \frac{d}{dt} (\boldsymbol{\omega} \cdot \overleftrightarrow{\mathbf{I}} \cdot \boldsymbol{\omega})
\end{aligned} \tag{10.30}$$

But, $T = \frac{1}{2}\boldsymbol{\omega} \cdot \overleftrightarrow{\mathbf{I}} \cdot \boldsymbol{\omega}$, by equation (10.19). Thus, we have

$$\begin{aligned}
\boldsymbol{\omega} \cdot \mathbf{N} &= \frac{1}{2} \frac{d}{dt} (\boldsymbol{\omega} \cdot \overleftrightarrow{\mathbf{I}} \cdot \boldsymbol{\omega}) \\
&= \frac{dT}{dt}
\end{aligned} \tag{10.31}$$

It should be noted that the differentiations in equations (10.30) and (10.31) can be carried out either in the fixed or in the body frame of reference. Equation (10.31) gives the relation between the rate at which work is done by the torque and the rate of change of kinetic energy with respect to time.

10.5 TORQUE-FREE MOTION

In order to solve Euler's equations of motion of a rigid body, the components of torque $\mathbf{N}$ along the principal axes of the rotating body must be known. However, this is seldom the case. The motion of a free symmetric top is the simplest type of the motion of a rigid body in which

the torque acting on it is known to be zero. As mentioned earlier, a body is called a free symmetric top if $I_1 = I_2 \neq I_3$.

Under these conditions, equations (10.27) become

$$I_1\dot{\omega}_1 = (I_1 - I_3)\omega_2\omega_3 \tag{10.32}$$

$$I_2\dot{\omega}_2 = (I_1 - I_3)\omega_3\omega_1 \tag{10.33}$$

and

$$I_3\dot{\omega}_3 = 0 \tag{10.34}$$

Equation (10.34) shows that ω_3, the component of the angular velocity along the z-axis is constant. Thus

$$\omega_3 = \text{const} \tag{10.35}$$

Equations (10.32) and (10.33) can be written as

$$\dot{\omega}_1 = \left[\frac{(I_1 - I_3)\omega_3}{I_1}\right]\omega_2$$

i.e.

$$\dot{\omega}_1 = \Omega\omega_2 \tag{10.36a}$$

and

$$\dot{\omega}_2 = -\Omega\omega_1 \tag{10.36b}$$

where

$$\Omega = \frac{I_1 - I_3}{I_1}\,\omega_3 = \text{const} \tag{10.37}$$

Differentiating equation (10.36a) with respect to time and substituting the value of $\dot{\omega}_2$ in it as obtained from equation (10.36b), we get

$$\ddot{\omega}_1 = -\Omega^2\omega_1 \tag{10.38}$$

Equation (10.38) resembles in form the differential equation of the simple harmonic motion. The solution of equation (10.38) is

$$\omega_1 = A\sin(\Omega t + \theta_0) \tag{10.39}$$

where A and θ_0 are the constants. Substituting this value of ω_1 in equation (10.36b) and integrating it, we get

$$\omega_2 = A\cos(\Omega t + \theta_0) \tag{10.40}$$

Thus, components ω_1 and ω_2 of the angular velocity change in such a manner that their resultant $\boldsymbol{\omega}_p$ is in the x-y plane and it rotates with angular frequency Ω. It is obvious that

$$\boldsymbol{\omega}_p = \mathbf{i}\omega_1 + \mathbf{j}\omega_2$$

The sense of Ω is the same as that of ω_3 if $I_1 > I_3$ and opposite to that of ω_3 if $I_1 < I_3$.

The magnitude of this vector is given by

$$|\boldsymbol{\omega}_p| = A$$

Further, since ω_3 is constant, the magnitude of the angular velocity vector $\boldsymbol{\omega}$ is also a constant. Thus

$$|\omega| = \sqrt{\omega_1^2 + \omega_2^2 + \omega_3^2} = \sqrt{\omega_p^2 + \omega_3^2} = \sqrt{A^2 + \omega_3^2} = \text{const.} \tag{10.41}$$

Thus, the angular velocity vector $\boldsymbol{\omega}$ rotates about the body z axis describing a cone with the vertex at the origin (Fig. 10.3). This motion is called *precession* and the body is said to precess about the z-axis with precessional velocity Ω. The cone described by the angular velocity vector $\boldsymbol{\omega}$ is known as the body cone, and its half angle α_b is given by

The axis of rotation of the earth is inclined at an angle of about 23.5° to the ecliptic plane of its orbit around the sun. This, together with the equatorial bulge gives rise to a torque due to the gravitational attractions exerted by the sun and the moon. The torque is very weak and produces a precessional motion, the period of which is approximately 2,600 years. This is the so-called *precession of equinoxes.*

In the torque-free motion of an asymmetric top, let us consider an interesting case, when the angular velocity of the body is nearly equal to ω_3 and components ω_1 and ω_2 are very small. The equations of motion are

$$I_1\dot{\omega}_1 = (I_2 - I_3)\omega_2\omega_3 \tag{10.49}$$

$$I_2\dot{\omega}_2 = (I_3 - I_1)\omega_3\omega_1 \tag{10.50}$$

$$I_3\dot{\omega}_3 = (I_1 - I_2)\omega_1\omega_2 = 0 \tag{10.51}$$

since both ω_1 and ω_2 are small. Thus, ω_3 is a constant upto the first order of ω_1 and ω_2. Differentiating equation (10.49) with respect to time and substituting for $\dot{\omega}_2$ from equation (10.50), we get

$$\ddot{\omega}_1 = -\frac{(I_2 - I_3)(I_1 - I_3)}{I_1I_2}\,\omega_3^2\omega_1 = -\Omega_1^2\omega_1 \tag{10.52}$$

where
$$\Omega_1 = \omega_3\sqrt{\frac{(I_2 - I_3)(I_1 - I_3)}{I_1I_2}} \tag{10.53}$$

This gives

$$\omega_1 = A\sin(\Omega_1 t + \theta_0) \tag{10.54}$$

where A and θ_0 are constants.

Substituting this value in equation (10.50) and simplifying, we get

$$\omega_2 = A\sqrt{\frac{I_1(I_1 - I_3)}{I_2(I_2 - I_3)}}\cos(\Omega_1 t + \theta_0) \tag{10.55}$$

The solutions expressed in equations (10.54) and (10.55) are valid as long as Ω_1 is real. Now, Ω_1 is real if I_3 is less than both I_1 and I_2, or I_3 is greater than both I_1 and I_2. In other words, Ω_1 is real if the moment of inertia about an axis along which ω is directed, has a maximum or a minimum value. If I_3 lies between I_1 and I_2, Ω_1 is imaginary. In that case, let $\Omega_1 = i\Omega_2$, where Ω_2 is a real quantity. Then, equation of motion, viz., equation (10.52) becomes

$$\ddot{\omega}_1 = \Omega_2^2\omega_1 \tag{10.56}$$

It has a general solution

$$\omega_1 = A_1e^{\Omega_2 t} + A_2e^{-\Omega_2 t} \tag{10.57}$$

Similarly, we can obtain

$$\omega_2 = B_1e^{\Omega_2 t} + B_2e^{-\Omega_2 t} \tag{10.58}$$

These solutions show that as time progresses, components ω_1 and ω_2 will increase. Thus, the initial assumptions no longer hold good and further analysis becomes inappropriate.

Thus, the torque-free rotational motion of a rigid body, having its

rotating body can perform three types of motion—a spin motion, a precessional motion and nutational motion.

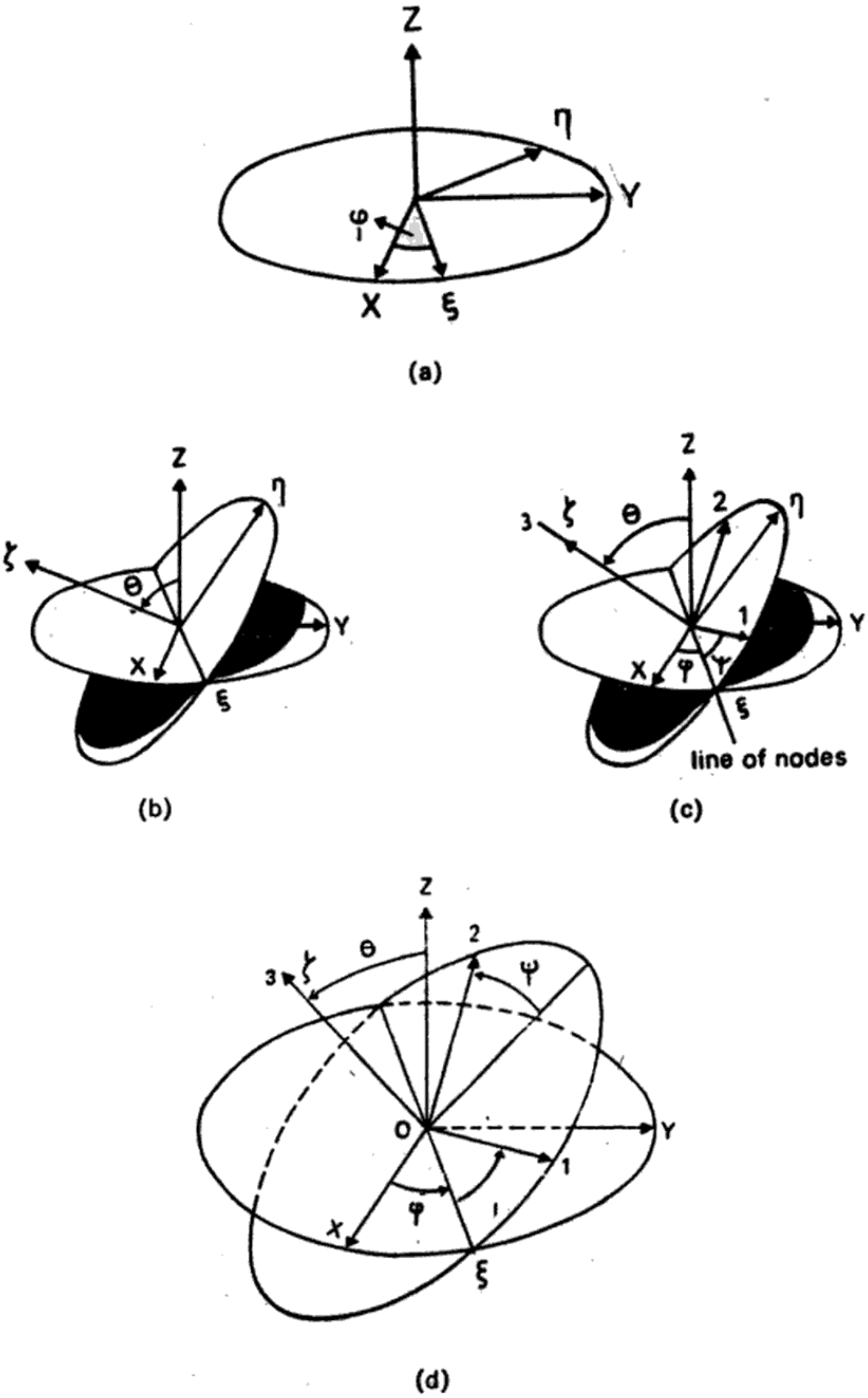


Fig. 10.5 Euler's angles: (a) rotation φ; (b) rotation θ; (c) rotation ψ; and (d) rotation from $O(XYZ)$ to $O(123)$

Let the angular velocity of the body be $\boldsymbol{\omega}$ which has components ω_1, ω_2 and ω_3 about the principal axes of the body frame of reference. We

can also resolve $\boldsymbol{\omega}$ along the ξ, η, ζ axes. Axes (123) rotate with an angular velocity $\dot{\psi}$ about axis 3 relative to the $\xi\eta\zeta$-axes. Hence, from Fig. 10.6, we can write

$$\omega_\xi = \dot{\theta},\ \omega_\eta = \dot{\varphi} \sin\theta \text{ and } \omega_\zeta = \dot{\psi} + \dot{\varphi}\cos\theta \tag{10.59}$$

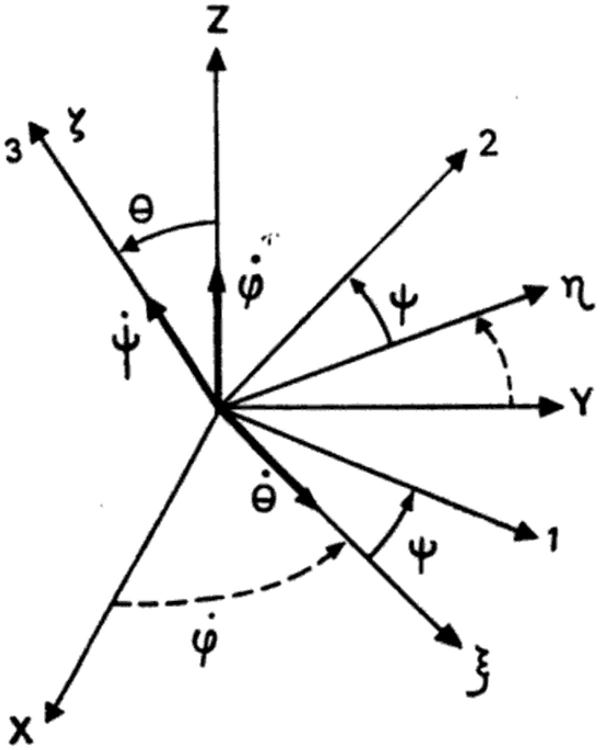


Fig. 10.6 Angular velocities $\dot{\varphi}$, $\dot{\psi}$ and $\dot{\theta}$

These components can be further resolved along the 1, 2, 3 axes and we get

$$\left.\begin{aligned} \omega_1 &= \omega_\xi \cos\psi + \omega_\eta \sin\psi = \dot{\theta}\cos\psi + \dot{\varphi}\sin\theta\sin\psi \\ \omega_2 &= -\omega_\xi \sin\psi + \omega_\eta \cos\psi = -\dot{\theta}\sin\psi + \dot{\varphi}\sin\theta\cos\psi \\ \omega_3 &= \omega_\zeta = \dot{\psi} + \dot{\varphi}\cos\theta \end{aligned}\right\} \tag{10.60}$$

The kinetic energy of the body is

$$T = \tfrac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2) \tag{10.61}$$

It is obvious that kinetic energy T will depend upon angular velocities $\dot{\theta}$, $\dot{\varphi}$ and $\dot{\psi}$ and coordinates θ and ψ. Coordinate φ is absent in the kinetic energy term and it will be an ignorable coordinate if it is absent in the potential energy also. The expression for kinetic energy T is a lengthy one and it will be considerably simplified if there exists an axis of symmetry for a rigid body. For a symmetric top $I_1 = I_2$ and the kinetic energy becomes

$$T = \tfrac{1}{2}I_1(\dot{\theta}^2 + \dot{\varphi}^2 \sin^2\theta) + \tfrac{1}{2}I_3(\dot{\psi} + \dot{\varphi}\cos\theta)^2 \tag{10.62}$$

Equation (10.62) shows that, for a symmetric top, kinetic energy T depends upon only one generalised coordinate θ. The potential energy, however, will, in general, depend upon all the three coordinates, i.e., $V \equiv V(\theta, \varphi, \psi)$. We can then write down Lagrange's equations of motion

tant momenta p_φ and p_ψ. This is because

$$\frac{p_\varphi - p_\psi \cos\theta}{I_1 \sin^2\theta} = \dot{\varphi} \tag{10.67}$$

Thus, we have

$$E = \tfrac{1}{2}I_1\dot{\theta}^2 + \frac{(p_\varphi - p_\psi \cos\theta)^2}{2I_1 \sin^2\theta} + \frac{p_\psi^2}{2I_3} + mgl\cos\theta \tag{10.66b}$$

Since the term $\frac{p_\psi^2}{2I_3}$ on the right hand side of equation (10.66b) is constant, we introduce another constant E' defined by

$$E' = E - \frac{p_\psi^2}{2I_3} = \tfrac{1}{2}I_1\dot{\theta}^2 + V(\theta) \tag{10.68}$$

where
$$V(\theta) = \frac{(p_\varphi - p_\psi \cos\theta)^2}{2I_1 \sin^2\theta} + mgl\cos\theta \tag{10.69}$$

is the effective potential energy. It depends upon θ alone.

The problem of motion of the symmetric top is solved in principle when we know the dependence of θ on time.
From equation (10.68), we have

$$\dot{\theta} = \frac{d\theta}{dt} = \sqrt{\frac{2}{I_1}[E' - V(\theta)]} \tag{10.70}$$

The solution of equation (10.70) is

$$t(\theta) = \int \frac{d\theta}{\sqrt{\frac{2}{I_1}[E' - V(\theta)]}} \tag{10.71}$$

Knowing $t(\theta)$ and hence $\theta(t)$, we can obtain the values of $\varphi(t)$ and $\psi(t)$ from equations (10.64) and (10.65). This method of solution involves elliptic integrals and is out of the scope of this book. We shall, therefore, obtain the qualitative features of the motion without performing these integrations. For this, we use the method of energy considerations used earlier in the central force-field motion.

Consider the variation of effective potential energy $V(\theta)$ with θ. The effective potential energy ranges between its values at $\theta = 0$ and $\theta = \pi$ and tends to infinity at the end values, as seen from equation (10.69).

Now
$$\frac{\partial V(\theta)}{\partial \theta} = \frac{(p_\varphi - p_\psi \cos\theta)(p_\psi - p_\varphi \cos\theta)}{I_1 \sin^3\theta} - mgl\sin\theta \tag{10.72}$$

In general, for $p_\varphi \neq p_\psi$, the quantity $\frac{\partial V(\theta)}{\partial t}$ is (i) positive for $\theta \to \pi$, and (ii) negative for $\theta \to 0$. Moreover, it has one zero value between 0 and π.

The variation of potential energy $V(\theta)$ with θ is shown in Fig. 10.8. The value of angle θ_0, when the potential energy $V(\theta)$ is minimum, is given by equation (10.72) when $\frac{\partial V(\theta)}{\partial \theta} = 0$.

Thus

$$(p_\varphi - p_\psi \cos\theta)(p_\psi - p_\varphi \cos\theta) = mglI_1 \sin^4\theta \tag{10.73}$$

is to be a 'precession without nutation'. The equality in condition (10.78) gives the minimum spin angular velocity at which the top will be just able to perform precession without nutation. If the spin angular velocity is below this value, the top will not be able to perform uniform precessional motion.

For a given value of θ_0, there are two values of precessional velocity $\dot{\varphi}$ given by equation (10.76). Thus, for a given spin angular velocity of the top thrown at angle θ_0, the top will start performing precessional motion with two possible velocities in the same sense—one is greater than the other. When velocity ω_3 and hence p_ψ is greater, i.e., in the case of a 'fast' top, the term under the radical sign can be expanded to give

$$\dot{\varphi}_{0_{\text{fast}}} = \frac{p_\psi}{2I_1 \cos\theta_0} = \frac{I_3\omega_3}{2I_1 \cos\theta_0} \tag{10.79}$$

Similarly, for the slow precession,

$$\dot{\varphi}_{0_{\text{slow}}} = \frac{mgl}{p_\psi} = \frac{mgl}{I_3\omega_3} \tag{10.80}$$

It is normally the slow precession that is observed in a rapidly spinning top.

Nutational Motion

It is seen above that the axis of the top oscillates between two limiting angles θ_1 and θ_2 for a given energy E'. These angles are given by

$$E' = V(\theta) = \frac{(p_\varphi - p_\psi \cos\theta)^2}{2I_1 \sin^2\theta} + mgl\cos\theta \tag{10.81}$$

This equation is a cubic in $\cos\theta$ and will have three roots. The potential energy curve (Fig. 10.8) indicates that only two out of the three roots lie in the physically possible range of $\cos\theta$, viz. -1 to $+1$ and the third must be trivial. When nutational motion is present the precessional velocity is given by equation (10.74), i.e.

$$\dot{\varphi} = \frac{p_\varphi - p_\psi \cos\theta}{I_1 \sin^2\theta} \tag{10.82}$$

It is clear that precessional velocity $\dot{\varphi}$ varies as θ changes. Let us define angle θ' such that

$$\cos\theta' = \frac{p_\varphi}{p_\psi} \tag{10.83}$$

for $p_\varphi < p_\psi$. Then, the precessional velocity is given by

$$\dot{\varphi} = \frac{p_\psi}{I_1 \sin^2\theta}(\cos\theta' - \cos\theta) \tag{10.84}$$

For any $\theta > \theta'$, $\cos\theta' > \cos\theta$ and $\dot{\varphi}$ has the sign of p_ψ and hence that of ω_3, since $p_\psi = I_3\omega_3$. If, however, $\theta < \theta'$, $\dot{\varphi}$ has a sign opposite to that of ω_3.

The derivative of the right-hand side of equation (10.81) with respect to time as seen from equation (10.72) is negative at $\theta = \theta'$. Hence, θ' corresponds to the angle less than θ_0. Thus, for the two angles θ_1 and θ_2, if $\theta_1 > \theta'$, then $\dot{\varphi}$ has the same sign as ω_3 throughout the nutational motion, whereas if $\theta_1 < \theta' < \theta_2$, $\dot{\varphi}$ has opposite directions as shown in Fig. 10.9. Figures 10.9a and 10.9b show the loci of the axis of the top on a unit sphere.

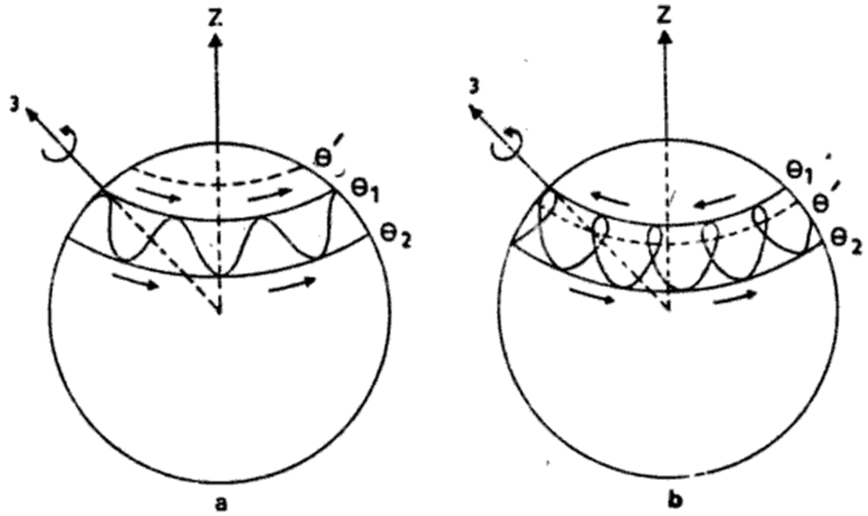


Fig. 10.9 Direction of precessional velocity for different values of θ
(a) $\theta_1 > \theta'$, (b) $\theta_1 < \theta' < \theta$

Consider now the case of a symmetric top started with a spin angular velocity $\omega_3 = \dot{\psi}$ at a given angle $\theta = \theta_1$. It is clear that at the beginning of the motion $\dot{\theta} = 0$ and $\dot{\varphi} = 0$. The constants of the motion in this case are obtained from equations (10.64) to (10.67).
These are

$$\left.\begin{aligned} p_\psi &= I_3\omega_3 \\ p_\varphi &= I_3\omega_3 \cos\theta_1 \\ E' &= mgl\cos\theta_1 \end{aligned}\right\} \qquad (10.85)$$

and

In this case, $\theta' = \theta_1$ and the motion of the top is shown in Fig. 10.10. The effective potential energy in this case is given by

$$V(\theta) = \frac{I_3^2\omega_3^2}{2I_1}\left[\frac{(\cos\theta_1 - \cos\theta)^2}{\sin^2\theta} + a\cos\theta\right] \qquad (10.86)$$

where $a = \dfrac{2mglI_1}{I_3^2\omega_3^2}$

Starting angle θ_1 is obviously one of the roots of equation (10.81) which can be written as

$$(\cos\theta_1 - \cos\theta)^2 - a(\cos\theta_1 - \cos\theta)(1 - \cos^2\theta) = 0 \qquad (10.87)$$

Equation (10.87) is cubic in $\cos\theta$ and has three roots. On factorising out one root $\cos\theta = \cos\theta_1$, we have

$$a\cos^2\theta - \cos\theta + \cos\theta_1 - a = 0$$

Then, the other two roots are

$$\cos \theta = \frac{1}{2a} [1 \pm \sqrt{1 - 4a \cos \theta_1 + a^2}] \tag{10.88}$$

Out of the two signs, the positive sign before the square root gives non-physical value for $\cos \theta$, viz. $\cos \theta > 1$, and the negative sign before the square root sign gives the value θ_2.

Thus, when the spinning top initially at rest is released at angle θ_1, it falls slightly under the action of gravity and gains in the precessional and the nutational motion. The axis of the top reaches the limiting angle θ_2 where it has maximum precessional velocity and zero nutational velocity. It describes successive cusps with the turning points on the surface at $\theta = \theta_1$ (Fig. 10.10). The precessional velocity at $\theta = \theta_2$ can be obtained from equation (10.82) and has the value

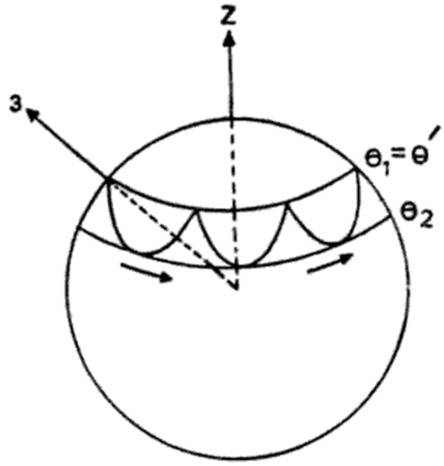


Fig. 10.10 Oscillation of the spin axis—nutation

$$\dot{\varphi} = \frac{I_3 \omega_3}{I_1} a = \frac{2mgl}{I_3 \omega_3} \tag{10.89}$$

Thus, the axis of the top undergoes precession very slowly (since $\dot{\varphi} = 0$) and nutation at a very large velocity at the smaller angle $\theta = \theta_1$. The situation is exactly opposite at the larger angle $\theta = \theta_2$. We have neglected, in all the above discussion, the frictional torque. If it is present, the nutational motion in most cases is completely damped and precession alone is observed uniformly.

QUESTIONS

1. What is meant by a rigid body? Write down the equation of constraint for such a body. Is it possible to have a perfectly rigid body? If not, why do we assume it to be rigid? Explain.
2. Distinguish between a symmetric top, spherical top and rotor.
3. Explain the terms: moments of inertia and products of inertia. When are the products of inertia zero? Explain with suitable illustrations.
4. What are Euler angles? Draw a neat diagram showing these angles. What will happen if their order is not maintained?
5. Why is moment of inertia not considered a vector nor a scalar?
6. Discuss the precessional motion. Give some illustrations of it.
7. State Newton's second law of motion as applied to a rotational motion.

how the remaining axes are chosen.

10. Show that a uniform sphere has vanishing products of inertia.
11. Show that the kinetic energy of a hoop rolling without slipping on a horizontal surface is $ma^2(\dot{\psi} + \dot{\varphi} \cos \theta)^2 + \frac{3}{4} ma^2\dot{\theta} - \frac{ma^2}{4} \sin^2 \theta\dot{\varphi}^2$. Obtain the equations of motion under gravity.
12. A gyroscope consists of a uniform circular disc of radius 4 cm and mass 10 g, with the centre of mass fixed and spinning at 2400 rpm. A 5g mass is attached to the axis 10 cm from the centre of mass. Find the rate of precession.
13. A cone of height 5 cm and base radius 2 cm, spins at 1200 rpm with its vertex fixed. Calculate I_3 and find the rate and period of precession of the axis about the vertical.
14. Investigate the motion of a symmetrical top in a gravitational field, one point on the axis of the top being held fixed. Show that the total energy E and the angular momenta p_{φ} and p_{ψ} about the vertical axis and about the symmetry axis of the top are constants of the motion.
15. A gyroscope consisting of a uniform solid sphere of radius 10 cm is spinning at 300 rpm about a horizontal axis. Due to faulty construction, the fixed point is not precisely at the centre of mass, but 0.02 mm away from it on the axis. Find the time taken by the axis to move through 1°.
16. A uniform rod of length l slides with its ends on a smooth vertical circle. If the length of the equivalent simple pendulum is r, find l/r.
17. A car starts from rest with a door open at a right angle. The door slams shut due to the acceleration, which is constant. Express the time required to close the door in terms of the relevant quantities. If the door is a uniform rectangle 1 m wide, and the acceleration is 0.5 m/s^2, find the time required to close the door.
18. Show that the angle of rotation φ is given in terms of the Euler angles by

$$\cos \frac{\varphi}{2} = \cos \frac{\varphi + \psi}{2} \cos \frac{\theta}{2}$$

19. Obtain the Lagrangian equations for the symmetric top.
20. Obtain the Lagrangian equations for the torque-free motion of a symmetric rigid body and calculate the motion of the axis of symmetry.
21. Calculate the moments of inertia I_1, I_2 and I_3 of a homogeneous sphere of radius r and mass M. Choose the origin of the axes at the centre of the sphere.
22. Determine the moments of inertia I_1, I_2 and I_3 of a homogeneous ellipsoid of mass M, given that the length of the axes are

$$2a > 2b > 2c$$

23. Show that none of the principal moments of inertia can exceed the sum of the other two.
24. Prove that the principal moments of inertia are all equal for any regular polyhedron whose centre is at the origin of the coordinate axes. Find the radius of a homogeneous sphere that has moments of inertia equal to those of a regular tetrahedron, the masses of the two being equal.
25. A top is made by forcing a light pin through the centre of a uniform disc whose radius is 3 cm. The pin projects 6 cm below the disc. The top is set into steady motion so that the rim of the disc just fails to touch the ground. Find the minimum number of revolutions made by the disc per second.

11

Variational Principle : Lagrange's and Hamilton's Equations

In this chapter we shall obtain Lagrange's equations from the variational principle. It will be recalled that Lagrange's equations were obtained in Chapter 8 from D'Alembert's principle in which Newton's law of motion, viz. $\mathbf{F} = m\mathbf{a}$ was used. D'Alembert's principle is often termed a 'differential principle' since we consider in it the instantaneous state of the system with some infinitesimal virtual displacements from that position of the system. The variational principle, in that sense, must be looked upon as an 'integral principle'. Here, we consider the entire motion of the system between two instants t_1 and t_2 and consider small virtual variations of the entire motion of the system from the actual motion.

11.1 CONFIGURATION SPACE

In the case of motion of a single particle we can represent its trajectory in the three-dimensional space by specifying its variables. For a system of N particles described by $3N$ space coordinates with k equations of constraint in the real space, it is difficult to visualise the motion of the entire system. It is, therefore, convenient to describe the state of a system having $3N - k = n$ coordinates in a hypothetical n-dimensional space. This is, in fact, an extension of the three-dimensional to the n-dimensional geometry. The instantaneous state of the system is then described by a point having generalised coordinates q_i, where $i = 1, 2, \ldots, n$. The point is called the system point and the n-dimensional space is known as the *configuration space*. At some later instant, the state of the system changes and it will be represented by some other

point in the configuration space. Thus, the system point moves in the configuration space tracing out a curve. This curve represents the path of motion of the entire system. Here, by the 'motion of the system' we mean the motion of the system point along this path in the configuration space. We can consider time as a parameter of the curve, since each point in the configuration space has one or more values of time associated with it. It should be clearly borne in mind that the n-dimensional configuration space does not have any bearing with the three-dimensional space which we can visualise physically. The path of the system point in this n-dimensional configuration space, therefore, is only a convenient way of visualising the actual motion of the N particles with n degrees of freedom.

11.2 SOME TECHNIQUES OF CALCULUS OF VARIATION

Before stating the variational principle, let us discuss some aspects of the calculus of variation needed for the development of theory in this chapter. The basic problem of calculus of variation is to find a path, $y = y(x)$, in one dimension between x_1 and x_2, such that the line integral of some function $f(y, y', x)$, where $y' = \dfrac{dy}{dx}$, is an extremum, i.e., maximum or minimum. The quantity f depends upon the functional form of the dependent variable $y(x)$ and is called a functional. Symbolically, the problem can be stated as: for a function $f(y, y', x)$, the integral

$$J = \int_{x_1}^{x_2} f(y, y', x)\, dx \tag{11.1}$$

along the path $y = y(x)$ between x_1 and x_2 is to be extremum.

Let (x_1, y_1) and (x_2, y_2) be two points in the space (Fig. 11.1). We have shown two varied paths between two extreme points $y(x_1) = y_1$ and $y(x_2) = y_2$.

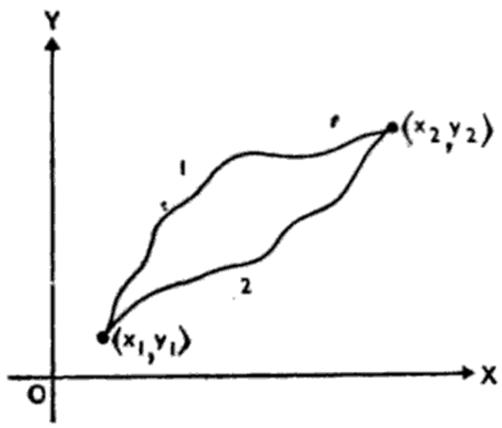


Fig. 11.1 Varied paths between two points

In order to find a path or paths which would give an extremum value for the integral, we state the problem in the language of differential calculus. For this, we associate a parameter α with all the possible curves i.e.

paths. The parameter α should be such that for some value of α, say $\alpha = 0$, the curve under examination would coincide with the path or paths that would give an extremum value for the integral. Then, y will be a function of both the independent variable x and the parameter α. We can always write $y(\alpha, x)$ as

$$y(\alpha, x) = y(0, x) + \alpha\eta(x) \tag{11.2}$$

where $\eta(x)$ is some function of x which has continuous first derivative and the function itself vanishes at both $x = x_1$ and $x = x_2$. The last condition, viz. $\eta(x_1) = \eta(x_2) = 0$ ensures that the varied function $y(\alpha, x)$ will be identical to $y(x)$ at the extremities of the path.

With the dependence of y on α in addition to x the integral becomes a function of the parameter α. Thus, we have

$$J(\alpha) = \int_{x_1}^{x_2} f[y(\alpha, x), y'(\alpha, x), x]\, dx \tag{11.3}$$

Then, the condition that $J(\alpha)$ has an extremum value is

$$\left[\frac{\partial J}{\partial \alpha}\right]_{\alpha=0} = 0 \tag{11.4}$$

This is only a necessary condition, but it is not sufficient. We shall, however, not go into the details of the mathematical aspect and use the condition to obtain a condition or equation that should be satisfied by f in order to have extremum value.

Differentiating equation (11.3) under the integral sign, we get

$$\frac{\partial J}{\partial \alpha} = \frac{\partial}{\partial \alpha}\left[\int_{x_1}^{x_2} f(y, y', x)\, dx\right]$$

$$= \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial \alpha}\right] dx$$

as $\dfrac{\partial x}{\partial \alpha} = 0$. Hence

$$\frac{\partial J}{\partial \alpha} = \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y'}\frac{\partial^2 y}{\partial \alpha\, \partial x}\right] dx \tag{11.5}$$

Integrating the second term in the integrand by parts, we get

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial y'}\frac{d}{dx}\left(\frac{\partial y}{\partial \alpha}\right) dx = \frac{\partial f}{\partial y'}\frac{\partial y}{\partial \alpha}\bigg|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\frac{\partial y}{\partial \alpha}\, dx \tag{11.6}$$

But

$$\frac{\partial y}{\partial \alpha} = \eta(x) \text{ and hence } \frac{\partial y}{\partial \alpha}\bigg|_{x_1}^{x_2} = \eta(x_2) - \eta(x_1) = 0 \tag{11.7}$$

Thus, the first term on the right-hand side of equation (11.6) vanishes and equation (11.5) becomes

$$\frac{\partial J}{\partial \alpha} = \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\frac{\partial y}{\partial \alpha}\right] dx$$

$$= \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] \eta(x)\, dx \tag{11.8}$$

It appears that equation (11.8) is independent of α. But, the functions y and y' with respect to which the derivatives of the functional f are taken, are functions of α. When $\alpha = 0$, we have $y(\alpha, x) = y(x)$, and the dependence on α disappears.

We want that $\left[\frac{\partial J}{\partial \alpha}\right]_{\alpha=0} = 0$, and since $\eta(x)$ is an arbitrary function [such that $\eta(x_1) = \eta(x_2) = 0$], the integrand of equation (11.8) must vanish for $\alpha = 0$. Thus, we have

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0 \tag{11.9}$$

Functions y and y' appearing in equation (11.9) are the original functions, independent of α. Equation (11.9) is called Euler's equation and it represents the necessary condition that the integral J has the extremum value.

Euler's equation can be generalised for the case when f is a functional of several dependent variables. Then, we can write

$$f = f[y_i(x), y_i'(x), x], \quad i = 1, 2, \ldots, n \tag{11.10}$$

In this case, we select the parametric equations similar to equation (11.2) as

$$y_i(\alpha, x) = y_i(0, x) + \alpha\eta_i(x) \tag{11.11}$$

The same procedure as described above can then be followed and we can get Euler-Lagrange's equation in the form

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\frac{\partial f}{\partial y_i'} = 0, \quad i = 1, 2, \ldots, n \tag{11.12a}$$

More generally

$$f = f[y_i(x_j), y_i'(x_j), x_j]$$

where $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, k$. Here, we have taken $x_1, x_2, \ldots, x_k$ as independent variables on which y depends. In this case, Euler-Lagrange's equations take the form

$$\frac{\partial f}{\partial y_i} - \sum_{j=1}^{k}\frac{\partial}{\partial x_j}\frac{\partial f}{(\partial y_i/\partial x_j)} = 0 \tag{11.12b}$$

Euler's equation (11.9) can also be put into another equivalent form. For this, consider

$$\begin{aligned}\frac{d}{dx}f(y, y', x) &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial x} \\ &= \frac{\partial f}{\partial x} + y'\frac{\partial f}{\partial y} + y''\frac{\partial f}{\partial y'}\end{aligned} \tag{11.13}$$

Now

$$\frac{d}{dx}\left(y'\frac{\partial f}{\partial y'}\right) = y''\frac{\partial f}{\partial y'} + y'\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)$$

Substituting the value of $y'' \frac{\partial f}{\partial y'}$ from equation (11.13), we get

$$\frac{d}{dx}\left(y' \frac{\partial f}{\partial y'}\right) = \frac{df}{dx} - \frac{\partial f}{\partial x} - y' \frac{\partial f}{\partial y} + y' \frac{d}{dx}\frac{\partial f}{\partial y'}$$

$$= \frac{df}{dx} - \frac{\partial f}{\partial x} - y'\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right] \qquad (11.14)$$

But, the last term on the right-hand side vanishes in view of equation (11.9) and equation (11.14) reduces to

$$\frac{\partial f}{\partial x} - \frac{d}{dx}\left(f - y' \frac{\partial f}{\partial y'}\right) = 0 \qquad (11.15)$$

This is sometimes called the 'second form' of Euler's equation. It is useful, particularly when $\frac{\partial f}{\partial x} = 0$, i.e., when f does not depend explicitly upon x. Then, we have

$$f - y' \frac{\partial f}{\partial y'} = \text{a const.} \qquad (11.16)$$

The δ-Notation

The results of the calculus of variation are very often expressed in terms of a compact δ-notation as follows:

We have, by equation (11.8)

$$\frac{\partial J}{\partial \alpha} = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right] \frac{\partial y}{\partial \alpha}\, dx$$

Multiply both the sides of this equation by the differential $d\alpha$. Then, we have

$$\frac{\partial J}{\partial \alpha}\, d\alpha = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right] \frac{\partial y}{\partial \alpha}\, d\alpha\, dx \qquad (11.17)$$

We introduce

$$\frac{\partial J}{\partial \alpha}\, d\alpha = \delta J \qquad (11.18)$$

and

$$\frac{\partial y}{\partial \alpha}\, d\alpha = \delta y$$

Hence, we get

$$\delta J = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right] \delta y\, dx \qquad (11.19)$$

In this notation, the condition of extremum becomes

$$\delta J = \delta \int_{x_1}^{x_2} f(y, y', x)\, dx = 0 \qquad (11.20)$$

Taking symbol δ inside the integral sign, we get

$$\delta J = \int_{x_1}^{x_2} \delta f\, dx$$

$$= \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y}\, \delta y + \frac{\partial f}{\partial y'}\, \delta y'\right] dx \qquad (11.21)$$

Now $$\delta y' = \delta\left(\frac{dy}{dx}\right) = \frac{d}{dx}(\delta y)$$

Hence $$\delta J = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y}\,\delta y + \frac{\partial f}{\partial y'}\frac{d}{dx}\,\delta y\right] dx \tag{11.22}$$

Integrating the second term on the right-hand side by parts, we get

$$\delta J = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right] \delta y\, dx \tag{11.23}$$

Here we used the condition that the variation in y at the end points is zero.

Since δy is arbitrary, we can get $\delta J = 0$ only if the integrand vanishes. This gives Euler-Lagrange's equation as before.

11.3 APPLICATIONS OF THE VARIATIONAL PRINCIPLE

1. Consider the extremum value of the integral

$$J = \int_{x_1}^{x_2} f(y, y', x)\, dx \tag{1}$$

where $f = \left(\frac{dy}{dx}\right)^2$ and $y(x) = x$

Clearly the value of the integral is $(x_2 - x_1)$. We construct a parametric equation as

$$y(\alpha, x) = x + \alpha\eta(x) \tag{2}$$

where we choose $\eta(x) = (x - x_1)(x - x_2)$, and satisfies the boundary conditions, i.e.

$$\eta(x_1) = \eta(x_2) = 0 \tag{3}$$

Now

$$\frac{dy}{dx} = 1 - \alpha(x_1 + x_2) + 2\alpha x = 1 - \alpha a - 2\alpha x$$

where $a = x_1 + x_2$.

Squaring and substituting the value, we get, after direct evaluation of integrals

$$J(\alpha) = (x_2 - x_1)(1 - 2\alpha a + \alpha^2 x^2) + \tfrac{4}{3}\alpha^2(x_2^3 - x_1^3) + 2\alpha(1 - \alpha a)(x_2^2 - x_1^2) \tag{4}$$

This shows that when $\alpha = 0$, $J = (x_2 - x_1)$. The value of the integral $J(\alpha)$ is always greater than $J(0)$ for any value of α. This fact can be stated as

$$\left.\frac{\partial J}{\partial \alpha}\right|_{\alpha=0} = 0 \tag{5}$$

2. To show that the shortest distance between two points in a plane is a straight line.

Consider the XY-plane. An element of length in this plane is given by

$$ds^2 = dx^2 + dy^2$$

or
$$ds = \sqrt{dx^2 + dy^2}$$
$$= dx\left[1 + \left(\frac{dy}{dx}\right)^2\right]^{1/2} \tag{1}$$

Then, the total distance between two points having coordinates (x_1, y_1) and (x_2, y_2) is given by
$$J = \int_1^2 ds = \int_{x_1}^{x_2} \left[1 + \left(\frac{dy}{dx}\right)^2\right]^{1/2} dx \tag{2}$$

The condition that the curve be the shortest path is that the integral J should be a minimum. Comparing this equation with equation (11.1) we obtain
$$f = \sqrt{1 + y'^2} \tag{3}$$
where $y' = \dfrac{dy}{dx}$.

Now
$$\frac{\partial f}{\partial y} = 0 \quad \text{and} \quad \frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1 + y'^2}}$$

Substituting these values in Euler-Lagrange's equation (11.9), we get
$$\frac{d}{dx}\left(\frac{y'}{\sqrt{1 + y'^2}}\right) = 0 \tag{4}$$
or $y'/\sqrt{1 + y'^2} = C$, where C is constant.

Simplifying this equation, we get
$$y' = \frac{C}{\sqrt{1 - C^2}} = a$$
where a is constant.

Integration, then gives
$$y = ax + b \tag{5}$$
where b is constant.

But this is the equation of a straight line. Thus, Euler-Lagrange's equation proves the familiar result.

3. The Brachistochrone or shortest time problem.

In this problem we consider a particle which moves in a constant conservative force-field $\mathbf{F}$ as shown in Fig. 11.2. Suppose that the particle is initially at rest at some point and moves to some other point (x_1, y_1) under the action of the force.

We wish to find the path that would be taken up by the particle along which its time of transit between the two points is a minimum. We orient the coordinate system such that the initial point at rest is taken as the origin.

Let v be the speed of the particle along the curve, the time of transit is given by
$$t_{12} = \int_1^2 \frac{ds}{v} \tag{1}$$
where 1 refers to the origin and 2 to the point (x_1, y_1).

If the frictional forces are ignored, we can write the total energy of the particle, i.e. $T + V =$ constant, since the force-field is given to be conservative. If the potential energy is taken to be zero at the level $x = 0$, we have $T + V = 0$ at the origin, since the particle is at rest there. Now, at any point (x, y) along its path the kinetic energy $T = \frac{1}{2}mv^2$ and the potential energy $V = -Fx = -mgx$, where g is the acceleration of the particle due to the force. Thus

$$v = \sqrt{2gx} \tag{2}$$

Hence, the time of transit is given by

$$t_{12} = \int_1^2 \frac{ds}{v} = \int_0^{x_1} \sqrt{(1 + y'^2)/2gx}\, dx \tag{3}$$

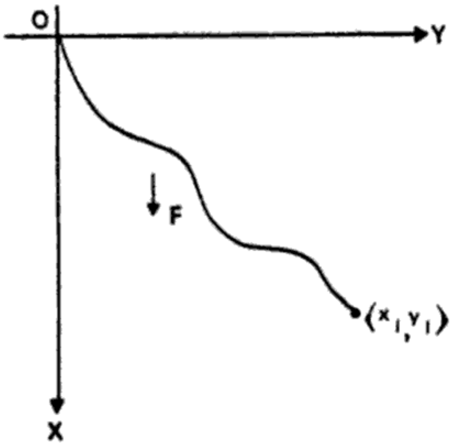


Fig. 11.2 The Brachistochrone problem

Since the factor $(2g)^{-1/2}$ does not affect the final equation, we can write functional f as

$$f = \sqrt{\frac{1 + y'^2}{x}} = \left[\frac{1 + y'^2}{x}\right]^{1/2} \tag{4}$$

As $\frac{\partial f}{\partial y} = 0$. Euler-Lagrange's equation becomes

$$\frac{d}{dx}\frac{\partial f}{\partial y'} = 0$$

or

$$\frac{\partial f}{\partial y'} = \text{const} = \frac{1}{\sqrt{2a}} \text{ by assumption}$$

Now

$$\frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{x(1 + y'^2)}} = \frac{1}{\sqrt{2a}}$$

Hence

$$\frac{y'^2}{x(1 + y'^2)} = \frac{1}{2a}$$

or

$$y' = \frac{\sqrt{x}}{\sqrt{2a - x}}$$

From this, we can always write

$$y = \int \frac{\sqrt{x}}{\sqrt{2a - x}} dx \tag{5}$$

To solve this, we substitute

$$x = a(1 - \cos\theta) \tag{6}$$

Hence,

$$dx = a \sin\theta \, d\theta \quad \text{and} \quad \sqrt{\frac{x}{2a - x}} = \tan\frac{\theta}{2}$$

Then, we get

$$y = \int a(1 - \cos\theta)\, d\theta$$
$$= a(\theta - \sin\theta) + \text{const.} \tag{7}$$

The constant term is equal to zero since initially when $x = 0 = y$, $\theta = 0$. Thus, we have the equations as

$$x = a(1 - \cos\theta)$$

and $$y = a(\theta - \sin\theta)$$

These are the equations of a cycloid passing through the origin. Thus, the path of the particle is a cycloid (Fig. 11.3). The value of constant a must be adjusted such that the particle passes through the other point, viz. (x_1, y_1). Along this path, the time of transit of the particle from the origin to (x_1, y_1) is found to be a minimum—although the variational principle is for an extremum value.

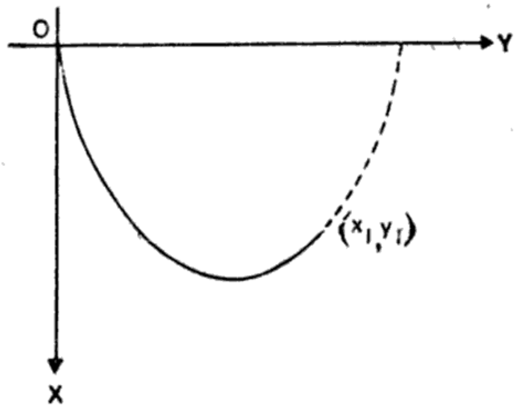


Fig. 11.3 The path of the particle is a cycloid

(4) We shall now show that the geodesics of a spherical surface are great circles, i.e., the circles whose centres lie at the centre of the sphere.

A 'geodesic' is a line which represents *the shortest path between any two points, when the path is restricted to be on some surface.* In the present case, the surface is a spherical surface.

The element of distance ds, on the surface of a sphere of radius r, in the spherical coordinates, is given by

$$ds^2 = r^2[d\theta^2 + \sin^2\theta \, d\varphi^2]$$

and $$ds = r\,[d\theta^2 + \sin^2\theta\, d\varphi^2]^{1/2} \tag{1}$$

The total distance between two points having coordinates (r, θ_1, φ_1) and (r, θ_2, φ_2) is given by

$$J = \int_1^2 ds = \int_{\theta_1}^{\theta_2} r\left[1 + \sin^2\theta\left(\frac{d\varphi}{d\theta}\right)^2\right]^{1/2} d\theta \tag{2}$$

Here, functional f is given by

$$f = r\left[1 + \sin^2\theta\left(\frac{d\varphi}{d\theta}\right)^2\right]^{1/2} \tag{3}$$

If J is to be extremum, we must have

$$\frac{\partial f}{\partial \varphi} - \frac{d}{d\theta}\frac{\partial f}{\partial \varphi'} = 0, \text{ where } \varphi' = \frac{d\varphi}{d\theta}$$

Since $\frac{\partial f}{\partial \varphi} = 0$, Euler-Lagrange's equation becomes

$$\frac{d}{d\theta}\frac{\partial f}{\partial \varphi'} = 0$$

or $$\frac{d}{d\theta}\left[\frac{\partial}{\partial \varphi'}\{r(1 + \sin^2\theta\,\varphi'^2)^{1/2}\}\right] = 0$$

This can be simplified to

$$\frac{d}{d\theta}\left[\frac{\sin^2\theta\,\varphi'}{(1 + \sin^2\theta\,\varphi'^2)^{1/2}}\right] = 0$$

since $r \neq 0$ and is constant.

From this, integration gives

$$\frac{\sin^2\theta\varphi'}{(1 + \sin^2\theta\varphi'^2)^{1/2}} = c \tag{4}$$

where c is constant.

Simplifying the above relation, we get

$$\varphi' = \frac{c}{\sin\theta(\sin^2\theta - c^2)^{1/2}} = \frac{c\,\text{cosec}^2\,\theta}{(1 - c^2 - c^2\cot^2\theta)^{1/2}} \tag{5}$$

Substitution for $\cot\theta$ leads to a standard form of integral for φ and integration yields

$$\varphi = \alpha - \sin^{-1}(k\cot\theta) \tag{6}$$

where α and $k = \frac{c}{\sqrt{1 - c^2}}$ are constants.

This gives

$$k\cot\theta = \sin(\alpha - \varphi)$$

or $$k\cos\theta = \sin(\alpha - \varphi)\sin\theta \tag{7}$$

Using the relations between the cartesian coordinates (x, y, z) and the spherical polar coordinates (r, θ, φ), we can write the above equation as

$$zk = r\sin\theta\cos\varphi\sin\alpha - r\sin\theta\sin\varphi\cos\alpha$$

or $$zk = x\sin\alpha - y\cos\alpha \tag{8}$$

where $x^2 + y^2 + z^2 = r^2$.

This last equation represents a plane passing through the origin and

Comparing equation (11.26) with equation (11.20) we obtain Euler-Lagrange's equations (11.12a), viz.

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\frac{\partial f}{\partial y_i'} = 0,\ i = 1, 2, \ldots, n$$

become

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0,\ i = 1, 2, \ldots, n$$

or

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0,\ i = 1, 2, \ldots, n \tag{11.27}$$

It will be recalled that equations (11.27) are Lagrange's equations of motion of a system of particles and the quantity $L = T - V$ is called Lagrange's function or the Lagrangian for the system. In terms of the Lagrangian L, Hamilton's principle can be stated as:

Of all the possible paths, along which a dynamical system may move from one point to another in the configuration space within a given interval of time, the actual path followed is that for which the time integral of the Lagrangian function for the system is an extremum.

Although the suffix i of equation (11.27) is stated to take all values from 1 to n, it should be recalled that this will not be the case if some constraints are acting on the system. In that case the number of independent coordinates (i.e., the degrees of freedom) is decreased and the number of Lagrange's equations is also correspondingly reduced.

11.5 EQUIVALENCE OF LAGRANGE'S AND NEWTON'S EQUATIONS

We have formulated mechanics on the basis of D'Alembert's principle in Chapter 8 and on Hamilton's principle in this chapter. We shall now show that the Newtonian formulation of mechanics is equivalent to the Lagrangian formulation by obtaining Newton's law of motion from Lagrange's equation, and Hamilton's principle from Newton's equation.

1. Let us write Lagrange's equation for a single particle in rectangular coordinates x_i, i.e. (x_1, x_2, x_3) as

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} - \frac{\partial L}{\partial x_i} = 0 \quad \text{where} \quad i = 1, 2, 3 \tag{11.28}$$

But, $L = T - V$. Hence, we get

$$\frac{d}{dt}\frac{\partial (T - V)}{\partial \dot{x}_i} - \frac{\partial (T - V)}{\partial x_i} = 0 \tag{11.29}$$

Now, in rectangular coordinates and for a conservative system, the kinetic energy T is a function of $\dot{x}_i$ alone and potential energy V is a function of x_i alone. Hence

$$\frac{\partial T}{\partial x_i} = 0 \quad \text{and} \quad \frac{\partial V}{\partial \dot{x}_i} = 0 \tag{11.30}$$

Lagrange's equation, therefore, assumes the form

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{x}_i} + \frac{\partial V}{\partial x_i} = 0 \tag{11.31}$$

or

$$-\frac{\partial V}{\partial x_i} = \frac{d}{dt}\frac{\partial T}{\partial \dot{x}_i} \tag{11.32}$$

For a conservative system, we have

$$-\frac{\partial V}{\partial x_i} = F_i \tag{11.33}$$

Further,

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{x}_i} = \frac{d}{dt}\frac{\partial}{\partial \dot{x}_i}\left(\sum_{j=1}^{3} \tfrac{1}{2} m \dot{x}_j^2\right)$$

$$= \frac{d}{dt}(m\dot{x}_i) = \frac{d}{dt}(p_i) = \dot{p}_i \tag{11.34}$$

Thus, we arrive at Newton's equation of motion, viz.

$$F_i = \dot{p}_i \tag{11.35}$$

The Lagrangian and Newtonian equations are, therefore, equivalent.

2. We now prove that Hamilton's principle may be obtained from Newton's equation. Consider the case of a single particle for simplicity. Let $x_i(t)$, $i = 1, 2, 3$ or $\mathbf{r}(t)$ be the solution of Newton's equation. Let $\mathbf{r}_1(t_1)$ and $\mathbf{r}_2(t_2)$ be the position vectors representing the position of the particle at instants t_1 and t_2. Then, Newton's equations and the equations of the constraint (if any) are satisfied at every point along the path of the particle.

Consider another path having the same end points and travelled in the same interval of time, viz. $(t_2 - t_1)$. Then, such a path would be represented by

$$\mathbf{r}(t) \to \mathbf{r}(t) + \delta\mathbf{r}(t) \tag{11.36}$$

Since the end points are the same for both the paths

$$\delta\mathbf{r}(t_1) = \delta\mathbf{r}(t_2) = 0 \tag{11.37}$$

Newton's equation of motion at any instant is

$$\mathbf{F} = m\mathbf{a} = m\ddot{\mathbf{r}} \tag{11.38}$$

Let δW be the work done in passing from the true path to the varied path. Then

$$\delta W = \mathbf{F}\cdot\delta\mathbf{r} \tag{11.39}$$

or

$$\delta W = m\ddot{\mathbf{r}}\cdot\delta\mathbf{r} \tag{11.40}$$

The total force $\mathbf{F}$ acting on the particle is the vector sum of the applied force $\mathbf{F}_a$ and the force of constraint $\mathbf{F}_c$. Thus, $\mathbf{F}_a + \mathbf{F}_c = \mathbf{F}$. The varied path considered above is such that no work is done by the force of constraint. Thus, the force of constraint and the displacement considered to obtain variation in the path are at right angles to each other. For example, when a particle moves on, say, a plane surface, the force of constraint is normal to the surface while the particle can move only along the surface. The varied path is obtained by considering displacements

parallel to the path along the surface. In such a case

$$\mathbf{F}_c \cdot \delta \mathbf{r} = 0$$

and equation (11.39) becomes

$$\delta W = \mathbf{F}_a \cdot \delta \mathbf{r} \tag{11.41}$$

If the applied force $\mathbf{F}_a$ is a conservative force and hence is derivable from a potential energy function V, then

$$\mathbf{F}_a \cdot \delta \mathbf{r} = -\delta V \tag{11.42}$$

Hence, equation (11.40) becomes

$$-\delta V = m\ddot{\mathbf{r}} \ \delta \mathbf{r} \tag{11.43}$$

Now, consider

$$\frac{d}{dt}(\dot{\mathbf{r}} \cdot \delta \mathbf{r}) = \dot{\mathbf{r}} \cdot \frac{d}{dt}(\delta \mathbf{r}) + \ddot{\mathbf{r}} \cdot \delta \mathbf{r} \tag{11.44}$$

On interchanging the operations $\frac{d}{dt}$ and δ in the first term on the right-hand side of equation (11.44), we get

$$\begin{aligned} \frac{d}{dt}(\dot{\mathbf{r}} \cdot \delta \mathbf{r}) &= \dot{\mathbf{r}} \cdot \delta \dot{\mathbf{r}} + \ddot{\mathbf{r}} \cdot \delta \mathbf{r} \\ &= \delta(\tfrac{1}{2}v^2) + \ddot{\mathbf{r}} \cdot \delta \mathbf{r} \end{aligned} \tag{11.45}$$

where $\dot{\mathbf{r}} = \mathbf{v}$. Thus

$$\ddot{\mathbf{r}} \cdot \delta \mathbf{r} = \frac{d}{dt}(\dot{\mathbf{r}} \cdot \delta \mathbf{r}) - \delta(\tfrac{1}{2}v^2) \tag{11.46}$$

Multiplying this equation by m throughout and writing $\delta(\frac{1}{2}mv^2)$ as δT and $m\ddot{\mathbf{r}} \cdot \delta \mathbf{r} = -\delta V$, we get

$$\delta T - \delta V = m \frac{d}{dt}(\dot{\mathbf{r}} \cdot \delta \mathbf{r}) \tag{11.47}$$

Integrating this last result with respect to time between t_1 and t_2, we get

$$\begin{aligned} \int_{t_1}^{t_2} \delta(T - V)dt &= m \int_{t_1}^{t_2} \frac{d}{dt}(\dot{\mathbf{r}} \cdot \delta \mathbf{r})\, dt \\ &= m \int_{t_1}^{t_2} d(\dot{\mathbf{r}} \cdot \delta \mathbf{r}) \\ &= m \Big[\dot{\mathbf{r}} \cdot \delta \mathbf{r}\Big]_{t_1}^{t_2} \end{aligned} \tag{11.48}$$

But, $\delta \mathbf{r} = 0$ at the end points. Hence, we get

$$\int_{t_1}^{t_2} \delta(T - V)\, dt = 0$$

or

$$\delta \int_{t_1}^{t_2} (T - V)\, dt = 0 \tag{11.49}$$

since the variation indicated by operator δ does not affect the end points. Equation (11.49) can be put in the form

$$\delta \int_{t_1}^{t_2} L\, dt = 0 \tag{11.50}$$

which is Hamilton's principle.

of voltages across each element of the circuit. Thus, we have

$$L\frac{dI}{dt} + RI + \frac{1}{C}\int I\,dt = \mathcal{E}(t) \qquad (11.52)$$

or by using $I = \frac{dq}{dt}$, where q is the charge

$$L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{q}{C} = \mathcal{E}(t) \qquad (11.53)$$

This equation is similar in form to equation (6.52) for the forced oscillator, viz.,

$$m\frac{d^2x}{dt^2} + 2m\mu\frac{dx}{dt} + kx = F(t) \qquad (11.54)$$

and hence must have the same solutions. This similarity suggests to us the mechanical analogies of electrical quantities in this case. The inductance L corresponds to inertial mass m, the ohmic resistance R corresponds to dissipative coefficient μ and $1/C$, where C is the capacitance, to k, the force constant or elastic force. The charge q plays the role of the coordinate x and the e.m.f. $\mathcal{E}(t)$ of the external force $F(t)$. Thus, the charge q can be treated as a generalised coordinate.

Consider, further, a circuit containing L, R and C in parallel connected with e.m.f. $\mathcal{E}(t)$ (Fig. 11.4b). In this case, the potential difference across each element is the same; but the currents flowing through R, L and C add to give the total current at any time. The current through R is $\frac{U}{R}$, that through L is $\frac{1}{L}\int U\,dt$, and that through C is $C\frac{dU}{dt}$. Therefore,

$$\frac{U}{R} + \frac{1}{L}\int U\,dt + C\frac{dU}{dt} = I(t)$$

Differentiating each term with respect to time, we have

$$C\frac{d^2U}{dt^2} + \frac{1}{R}\frac{dU}{dt} + \frac{U}{L} = \frac{dI}{dt} \qquad (11.55)$$

Comparison of equation (11.55) with equation (11.54) shows that the voltage U corresponds to x, C to m, $\frac{1}{R}$ to $2\,m\mu$, $\frac{1}{C}$ to k and $\frac{dI}{dt}$ corresponds to external force $F(t)$.

With these analogies, we can write down expressions for the kinetic and the potential energies, the dissipation function and the generalised force for L, C and R in series as

$$T = \tfrac{1}{2}L\dot{q}^2,\ V = \frac{q^2}{2C},\ \mathcal{F} = \tfrac{1}{2}R\dot{q}^2,\ Q(t) = \mathcal{E}(t) \qquad (11.56)$$

and for L, C and R in parallel as

$$T = \tfrac{1}{2}C\dot{U}^2,\ V = \frac{U^2}{2L},\ \mathcal{F} = \frac{\dot{U}^2}{2R},\ Q(t) = \frac{dI}{dt} \qquad (11.57)$$

The Lagrangian $L = T - V$ for the two cases can be written and equations (11.53) and (11.55) can be obtained by using corresponding dissipation functions.

Thus, the Lagrangian and Hamilton's principle along with variational technique offer a wide possibility of their use in diverse fields of physics. One such application of this technique in writing Schrödinger's equation in quantum mechanics is given in article 11.9(d).

11.7 LAGRANGE'S UNDETERMINED MULTIPLIERS

If a physical system is constrained in its motion, then its degrees of freedom are reduced. We use the equations of constraint to eliminate dependent variables and we work with a new set of independent variables. Sometimes it is difficult or inconvenient to eliminate the dependent variables. Under these circumstances, use of Lagrange's multipliers gives an alternative technique to solve the problems.

Consider a simple case of a function $f = f(x, y, z)$ of three independent variables. The function f has an extremum value when

$$df = 0 \tag{11.58}$$

Since

$$df = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy + \frac{\partial f}{\partial z}\, dz \tag{11.59}$$

for the condition (11.58) to be true, the necessary and sufficient conditions are

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \frac{\partial f}{\partial z} = 0 \tag{11.60}$$

Let the equation of constraint be

$$g(x, y, z) = 0 \tag{11.61}$$

or

$$dg = \frac{\partial g}{\partial x}\, dx + \frac{\partial g}{\partial y}\, dy + \frac{\partial g}{\partial z}\, dz \tag{11.62}$$

Because of equation (11.61) of constraint, condition (11.60) is no longer valid since there are now only two independent variables. If they are x and y, then dz is no longer arbitrary but will be related to changes in x and y. We can, however, add equations (11.61) and equation (11.62) after multiplying the latter by λ to get

$$\begin{aligned} df + \lambda\, dg &= \left(\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y}\right) dy + \left(\frac{\partial f}{\partial z} + \lambda \frac{\partial g}{\partial z}\right) dz \\ &= 0 \end{aligned} \tag{11.63}$$

The multiplier λ can be chosen by setting

$$\frac{\partial f}{\partial z} + \lambda \frac{\partial g}{\partial z} = 0 \tag{11.64}$$

where we assume that $\dfrac{\partial g}{\partial z}$ is non-zero.

D'Alembert's or Hamilton's principle, the condition that the constraints must be holonomic does not appear until the last step. There, the variations in the generalised coordinates are considered independent of each other. When the constraints acting on a system are nonholonomic, the generalised coordinates are not independent of each other. Moreover, it may not be possible to reduce them further by means of equations of constraint of the form $f(q_1, q_2, \ldots, q_n; t) = 0$.

Let us suppose that the equations of constraint can be put in the form

$$\sum_k a_{lk}\, dq_k + a_{lt}\, dt = 0,\ l = 1, 2, \ldots, m \tag{11.72}$$

Equation (11.72) represents m relations of constraint between the differentials of q's. Since, Hamilton's principle does not involve variation in time, the virtual displacements must satisfy the equation

$$\sum_k a_{lk}\, \delta q_k = 0,\ l = 1, 2, \ldots, m \tag{11.73}$$

Equations (11.73) can be used to reduce the number of virtual displacements to only that of the independent ones. For this, we follow the method of Lagrange's undetermined multipliers.

Let λ_l $(l = 1, 2, \ldots, m)$ be some unknown constants, which may be, in general, functions of time. Then, in view of equation (11.73), we can write

$$\lambda_l \sum_k a_{lk}\, \delta q_k = 0 \tag{11.74}$$

The m equations expressed in (11.74) are now combined with

$$\int_{t_1}^{t_2} \left[\sum_k \left(\frac{\partial L}{\partial q_k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} \right) \delta q_k \right] dt = 0 \tag{11.75}$$

for the conservative system.

For this, we sum up equation (11.74) over l and then integrate the resulting equation from t_1 to t_2. This yields

$$\int_{t_1}^{t_2} \sum_{kl} \lambda_l a_{lk}\, \delta q_k\, dt = 0 \tag{11.76}$$

Now, we combine equation (11.76) with equation (11.75), so as to get

$$\int_{t_1}^{t_2} \left[\sum_{k=1}^{n} \left(\frac{\partial L}{\partial q_k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} + \sum_l \lambda_l a_{lk} \right) \delta q_k \right] dt = 0 \tag{11.77}$$

The δq_k's are still dependent upon each other. In fact, the m equations (11.73) are really speaking the links between them. Thus, the first $(n - m)$ of these δq_k's may be chosen independently. Then, the last m of these δq_k's will be given by equation (11.73). But, λ_l's can be chosen arbitrarily. We may choose them such that

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} + \sum_k \lambda_l a_{lk} = 0 \tag{11.78}$$

where $k = (n - m + 1), \ldots, n$.

If this is true, we must write equation (11.77) as

$$\int_{t_1}^{t_2} \left[\sum_{k=1}^{n-m} \left(\frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} + \sum_l \lambda_l a_{lk} \right) \delta q_k \right] dt = 0 \tag{11.79}$$

Equation (11.79) involves only those δq_k's that are independent. Hence, the quantity in parentheses must vanish. Thus, we have

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} + \sum_l \lambda_l a_{lk} = 0,\ k = 1, 2, \ldots, (n-m) \tag{11.80}$$

Combining equations (11.78) and (11.80), we get a complete set of Lagrange's equations for nonholonomic systems:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = \sum_l \lambda_l a_{lk},\ k = 1, 2, \ldots, n \tag{11.81}$$

In the process described above, there are $n + m$ quantities to be determined out of which there are n generalised coordinates q_k and m multipliers λ_l. We, however, have only n equations (11.81) at our disposal. We need m more equations which are equations (11.72), written in a differential form as

$$\sum_k a_{lk} \dot{q}_k + a_{lt} = 0 \tag{11.82}$$

where $l = 1, 2, \ldots, m$.

On solving these equations not only do we get the values of q_k's, but also the values of λ_l's. To understand the physical significance of the λ_l's, consider a system on which an external force Q'_k acts instead of the constraints. Thus, the extra forces Q'_k not included in the conservative forces are given by

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = Q'_k \tag{11.83}$$

These equations must be identical with equations (11.81). Hence, $Q'_k = \sum \lambda_l a_{lk}$. Thus, $\sum \lambda_l a_{lk}$ can be treated as generalised forces of constraint. Thus, we do not eliminate the forces of constraint from the formulation. On the contrary, these are found out.

We have considered a method to account for: nonholonomic constraints which can be put in the form of equation (11.72) and is not applicable for the most general type of nonholonomic constraint. The method, however, also considers holonomic constraints. An equation of the holonomic constraint, viz.

$$f(q_1, q_2, \ldots, q_n; t) = 0$$

is equivalent to a differential equation

$$\sum_k \frac{\partial f}{\partial q_k} dq_k + \frac{\partial f}{\partial t} dt = 0 \tag{11.84}$$

Comparing equation (11.84) with equation (11.72), we obtain

$$a_{lk} = \frac{\partial f}{\partial q_k} \quad \text{and} \quad a_{lt} = \frac{\partial f}{\partial t} \tag{11.85}$$

Hence, the Lagrangian is given by

$$L = T - V = \tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{4}mr^2\dot{\theta}^2 - mg(l - x)\sin\varphi \tag{5}$$

The Lagrangian equations in x and θ are, therefore, given by

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = a_x\lambda$$

or

$$m\ddot{x} - mg\sin\varphi + \lambda = 0 \tag{6}$$

and

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = a_\theta\lambda$$

or

$$\tfrac{1}{2}mr^2\ddot{\theta} - r\lambda = 0 \tag{7}$$

Now, from equation (1) of constraint, we get

$$r\ddot{\theta} = \ddot{x} \tag{8}$$

Hence, Lagrange's equation in θ becomes

$$\tfrac{1}{2}m\ddot{x} - \lambda = 0$$

or

$$\frac{m\ddot{x}}{2} = \lambda \tag{9}$$

Substituting this value of λ in Lagrange's equation for x, we get

$$m\ddot{x} - mg\sin\varphi + \tfrac{1}{2}m\ddot{x} = 0$$

or

$$\ddot{x} = \frac{2g\sin\varphi}{3} \tag{10}$$

Hence

$$\ddot{\theta} = \frac{2g\sin\varphi}{3r} \tag{11}$$

The frictional force of constraint λ is given by

$$\lambda = \frac{m\ddot{x}}{2} = \frac{mg\sin\varphi}{3} \tag{12}$$

If we write $\ddot{x} = v\dfrac{dv}{dx}$ and integrate the equation $\ddot{x} = v\dfrac{dv}{dx} = \dfrac{2g\sin\varphi}{3}$, we obtain the velocity $v = \sqrt{4gl\sin\varphi/3}$ at the bottom of the inclined plane. This last result can be obtained by any elementary method as well.

(b) Simple Pendulum

As the next illustration, let us consider the problem of a simple pendulum.

Let (r, θ) be the coordinates of the bob of the pendulum with respect to the point O of the support (Fig. 11.6). Then, the Lagrangian is given by

$$L = T - V = \tfrac{1}{2}m\dot{r}^2 + \tfrac{1}{2}mr^2\dot{\theta}^2 + mgr\cos\theta \tag{1}$$

The equation of constraint is

$$r - l = 0$$

or

$$dr = 0$$

Hence

$$a_r = 1 \tag{2a}$$

and

$$a_\theta = 0 \tag{2b}$$

Lagrange's equations in r and θ are expressed as

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = a_r\lambda$$

or $$\frac{d}{dt}(m\dot{r}) - mr\dot{\theta}^2 - mg\cos\theta = \lambda \qquad (3)$$

and gives tension in the suspension λ in terms of θ and $\dot{\theta}$. Further

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = a_\theta\lambda$$

or $$\frac{d}{dt}(mr^2\dot{\theta}) + mgr\sin\theta = 0 \qquad (4)$$

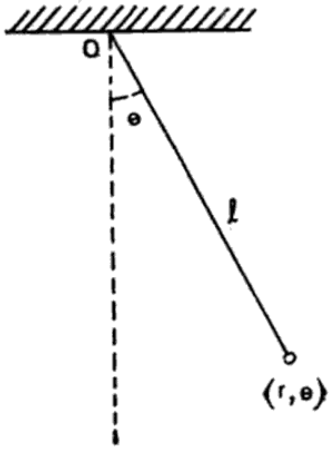


Fig. 11.6 A simple pendulum

This equations (4) gives

$$mr^2\ddot{\theta} + mgr\sin\theta = 0 \qquad (5)$$

which is the equation of motion of the simple pendulum.

(c) Particle on Sphere

Let us now consider a particle of mass m moving under the action of gravity on the surface of a smooth sphere of radius l.

We shall find its equations of motion and the angle θ_c at which the particle flies off from the surface.

Let the origin of the coordinates be at the centre of the sphere and let the z axis be vertically upwards.

In this case, the equation of constraint is given by

$$r - l = 0 \qquad (1)$$

where r is the radial distance of the particle.

From this, we obtain, $dr = 0$. Hence, $a_r = 1$, $a_\theta = 0 = a_\varphi$.

Let us suppose that the particle is initially at rest and let it slide down along the surface. The particle will obviously move in a vertical plane which we shall take for convenience $\varphi = 0$.

The Lagrangian for the particle is

$$L = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - mgr\cos\theta \qquad (2)$$

The equations of motion are

$$-ml\dot{\theta}^2 + mg\cos\theta = \lambda \qquad (3)$$

$$ml^2\ddot{\theta} - mgl\sin\theta = 0 \qquad (4)$$

where we used $r = l$ and $\dot{r} = \ddot{r} = 0$.

The undetermined multiplier λ is dependent on θ, in general.

Differentiating equation (3) containing λ with respect to time, we get

$$-2ml\dot{\theta}\ddot{\theta} - mg\sin\theta\,\dot{\theta} = \frac{d\lambda(\theta)}{dt} = \frac{d\lambda}{d\theta}\dot{\theta}$$

or $$-3mg\sin\theta = \frac{d\lambda}{d\theta} \qquad (5)$$

where we used value of $\ddot{\theta}$ from equation (4). Integrating equation (5), we get

$$\lambda(\theta) = 3mg \cos \theta + c$$

At $\theta = 0, \lambda = mg$, this being the force of constraint at the top of the sphere. This gives

$$c = -2mg$$

Hence $$\lambda(\theta) = 3mg \cos \theta - 2mg \tag{6}$$

The particle will move on the surface as long as the force of constraint is positive, i.e., as long as the surface pushes it outward. Corresponding condition is

$$\lambda(\theta) = 3mg \cos \theta - 2mg \geqslant 0 \tag{7}$$

The equality holds for

$$\cos \theta_c = \tfrac{2}{3} \tag{8}$$

i.e., at the angle $\theta_c = \cos^{-1} \frac{2}{3}$ the particle flies off the surface. It should be noted that we have neglected friction of the surface.

(d) The Schrödinger Wave Equation

The Lagrangian formulation is an important tool particularly in modern physics. We illustrate this by taking a quantum mechanical problem of variation in

$$\delta \int \psi^*(r) H(r, p) \psi(r)\, d\tau = 0 \tag{1}$$

with a constraint

$$\int \psi^*(r) \psi(r)\, d\tau = 1 \tag{2}$$

Equation (1) states that the energy of a particle described by wave function ψ is an extremum with the condition given by equation (2) that the total probability of finding the particle in the whole space is unity—a certainty. Here, H is the quantum mechanical Hamiltonian operator for a particle of mass m

$$H = -\frac{\hbar^2}{2m} \nabla^2 + V(r) \tag{3}$$

$V(r)$ being the potential field in which the particle is moving. We shall assume that the particle is confined in a finite potential field so that it is not found at a large distance and hence ψ and ψ^* vanish at large distances.

The wave function of the particle ψ and its complex conjugate are treated as independent variables.

Equation (1) when H from equation (3) is substituted, contains terms which after integration by parts become

$$\begin{aligned}\int \psi^* \frac{\partial^2 \psi}{\partial x^2}\, dx &= \psi^* \frac{\partial \psi}{\partial x}\bigg|_{-\infty}^{+\infty} - \int \frac{\partial \psi^*}{\partial x} \frac{\partial \psi}{\partial x}\, dx \\ &= -\int \frac{\partial \psi^*}{\partial x} \frac{\partial \psi}{\partial x}\, dx \end{aligned} \tag{4}$$

We have already introduced the Hamiltonian function which is related to the Lagrangian function by the equation

$$H = \sum p_i \dot{q}_i - L(q_i, \dot{q}_i, t)$$

Let the Hamiltonian be expressed as a function of generalised coordinates and generalised momenta $p_i = \dfrac{\partial L}{\partial \dot{q}_i}$, i.e., $\dot{q}_i$ is replaced by p_i in the above expression and

$$H \equiv H(q_i, p_i, t) \tag{11.86}$$

We want to describe the motion of the system in terms of an equation of motion involving the Hamiltonian. This clearly becomes a problem of transformation from the set of variables $(q_i, \dot{q}_i)$ to a new set (q_i, p_i). In order to achieve this transformation we write the differential form

$$dH = \sum_k \frac{\partial H}{\partial p_k} dp_k + \sum_k \frac{\partial H}{\partial q_k} dq_k + \frac{\partial H}{\partial t} dt \tag{11.87}$$

But, from the definition of H quoted above, we have

$$dH = \sum_k \dot{q}_k dp_k + \sum_k p_k d\dot{q}_k - dL \tag{11.88}$$

Since, $L \equiv L(q_i, \dot{q}_i, t)$, we get

$$dL = \sum_k \frac{\partial L}{\partial q_k} dq_k + \sum_k \frac{\partial L}{\partial \dot{q}_k} d\dot{q}_k + \frac{\partial L}{\partial t} dt$$

$$= \sum_k \frac{\partial L}{\partial q_k} dq_k + \sum_k p_k d\dot{q}_k + \frac{\partial L}{\partial t} dt \tag{11.89}$$

where we have used the definition of generalised momenta $p_k = \partial L/\partial \dot{q}_k$.

Substituting this value of dL in equation (11.88), we get

$$dH = \sum_k \dot{q}_k\, dp_k + \sum_k p_k\, d\dot{q}_k - \sum_k \frac{\partial L}{\partial q_k} dq_k - \sum_k p_k\, d\dot{q}_k - \frac{\partial L}{\partial t} dt$$

$$= \sum_k \dot{q}_k\, dp_k - \sum_k \dot{p}_k\, dq_k - \frac{\partial L}{\partial t} dt \tag{11.90}$$

Comparing the coefficients of dp_k and dq_k in equations (11.87) and (11.90), we get

$$\dot{q}_k = \frac{\partial H}{\partial p_k} \tag{11.91}$$

and

$$-\dot{p}_k = \frac{\partial H}{\partial q_k} \tag{11.92}$$

Equations (11.91) and (11.92) are called Hamilton's equations or Hamilton's canonical equations of motion. They form a set of $2n$ first-order differential equations of motion and replace the n-Lagrange equations of second order.

On comparing the coefficients of dt in equations (11.87) and (11.90), we get

$$-\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t} \tag{11.93}$$

If H does not involve a particular co-ordinate q_k, $\frac{\partial H}{\partial q_k} = 0$, i.e., $\dot{p}_k = 0$ or $p_k =$ constant. Such a coordinate q_k is called a *cylic* or *ignorable* coordinate as before. Thus, a cyclic coordinate in the Lagrangian will be absent in the Hamiltonian.

Since, $H \equiv H(p_k, q_k, t)$, the total time derivative of H is given by

$$\frac{dH}{dt} = \sum_k \frac{\partial H}{\partial q_k} \dot{q}_k + \sum_k \frac{\partial H}{\partial p_k} \dot{p}_k + \frac{\partial H}{\partial t} \tag{11.94}$$

Substituting Hamilton's equation in this equation, we find that

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \tag{11.95}$$

Hence

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \tag{11.96}$$

by equation (11.93).

If the Lagrangian does not involve time explicitly, $\frac{\partial L}{\partial t} = 0$

This gives

$$\frac{dH}{dt} = 0 \tag{11.97}$$

or

$$H = \text{const.} \tag{11.98}$$

Thus, the Hamiltonian H is a constant of the motion.

Now, for a conservative system, when the potential energy V is not a function of velocities, i.e., $\frac{\partial V}{\partial \dot{q}_k} = 0$, we have already shown in Chapter 8 that

$$\begin{aligned} \text{the Hamiltonian } H &= T + V \\ &= \text{kinetic energy} + \text{potential energy} \\ &= \text{total energy of the system} \end{aligned} \tag{11.99}$$

11.11 SOME APPLICATIONS OF THE HAMILTONIAN FORMULATION

Let us now illustrate the theory of Hamilton's equation with the help of the following examples.

(a) A Simple Pendulum with Moving Support

A pendulum of mass m is suspended from a support which is constrained to move along a straight horizontal line. The pendulum oscillates in the vertical plane containing the direction of motion of the support. We want to obtain the Hamiltonian of the system.

We direct the y-axis along the motion of the support Q and choose the angle θ with respect to the vertical to locate the pendulum at P (Fig. 11.7).

Let $PQ = l$. Then, the coordinates of P are

$$X = l\cos\theta \quad \text{and} \quad Y = y + l\sin\theta$$

Hence

$$\dot{X} = -l\dot{\theta}\sin\theta \quad \text{and} \quad \dot{Y} = \dot{y} + l\dot{\theta}\cos\theta$$

Now, the kinetic energy of the pendulum is

$$T = \tfrac{1}{2}mv^2 = \tfrac{1}{2}m(\dot{y}^2 + l^2\dot{\theta}^2 + 2\dot{y}l\dot{\theta}\cos\theta)$$

and its potential energy, with respect to its vertical position when $\theta = 0$, is

$$V = mgl(1 - \cos\theta)$$

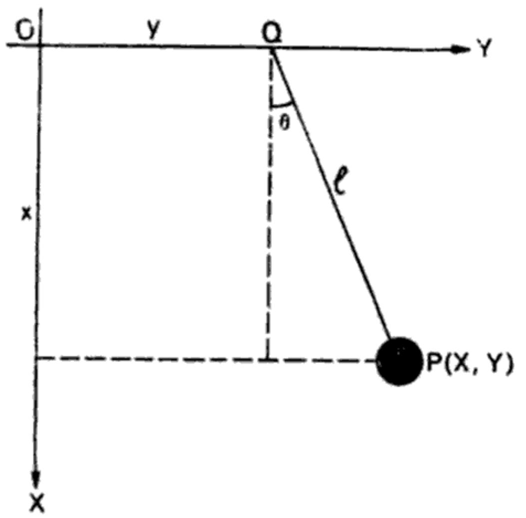


Fig. 11.7 A simple pendulum with a moving support

Hence, the Lagrangian is given by

$$L = T - V = \tfrac{1}{2}m(\dot{y}^2 + l^2\dot{\theta}^2 + 2\dot{y}l\dot{\theta}\cos\theta) - mgl(1 - \cos\theta) \tag{1}$$

The Hamiltonian function is

$$H = \sum_k p_k\dot{q}_k - L = p_y\dot{y} + p_\theta\theta - L \tag{2}$$

Here

$$p_y = \frac{\partial L}{\partial \dot{y}} = m(\dot{y} + l\dot{\theta}\cos\theta) \tag{3}$$

and

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = ml(l\dot{\theta} + \dot{y}\cos\theta) \tag{4}$$

or

$$\left.\begin{aligned} \dot{y} &= \frac{1}{m\sin^2\theta}\left(p_y - \frac{p_\theta}{l}\cos\theta\right) \\ l\dot{\theta} &= -\frac{1}{m\sin^2\theta}\left(p_y\cos\theta - \frac{p_\theta}{l}\right) \end{aligned}\right\} \tag{5}$$

and

$$\frac{1}{m}\left(p_y + \frac{p_\theta}{l}\right) = (\dot{y} + l\dot{\theta})(1 + \cos\theta) \tag{6}$$

From equation (2), we obtain the Hamiltonian as

$$H = m(\dot{y}^2 + \dot{y}l\dot{\theta}\cos\theta) + m(l^2\dot{\theta}^2 + \dot{y}l\dot{\theta}\cos\theta)$$
$$-\tfrac{1}{2}m(\dot{y}^2 + l^2\dot{\theta}^2 + 2\dot{y}l\dot{\theta}\cos\theta) + mgl(1 - \cos\theta)$$
$$H = \tfrac{1}{2}m(\dot{y}^2 + l^2\dot{\theta}^2 + 2\dot{y}l\dot{\theta}\cos\theta) + mgl(1 - \cos\theta) \tag{7}$$

Let us now consider a simple case of the motion of a charge q in a uniform magnetic field **B**. If we take **B** along the z-axis, then the vector potential given by $\mathbf{B} = \nabla \times \mathbf{A}$ has the value $B = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}$. A simple choice of components of vector potential in this case is

$$A_x = A_z = 0 \quad \text{and} \quad A_y = xB$$

The Hamiltonian in this case is

$$H = \frac{1}{2m}(p_x^2 + p_z^2) + \frac{1}{2m}(p_y - qxB)^2 \tag{7}$$

Since H does not depend on y and z, we have

$$p_y = \text{const. and } p_z = \text{const.}$$

By putting $\omega = \frac{qB}{m}$ and $x_0 = \frac{p_y}{qB}$ in equation (7), we get the Hamiltonian in the form

$$H = \frac{1}{2m}(p_x^2 + p_z^2) + \tfrac{1}{2}m\omega^2(x - x_0)^2 \tag{8}$$

We see here that x and p_x are obtained in H as in the harmonic oscillator. Hamilton's equations of motion are

$$\dot{p}_x = -m\omega^2(x - x_0),\ \dot{p}_y = 0 = \dot{p}_z \tag{9}$$

or $$p_y = \text{const. and } p_z = \text{const.}$$

The last two equations are already noted above from the absence of y and z in H. The first equation in equation (9) is the equation of simple harmonic oscillator

$$\ddot{x} = -\omega^2(x - x_0) \tag{10}$$

and has the following solution:

$$x = a\cos(\omega t + \alpha) + x_0 \tag{11}$$

It should be noted that x_0 is not a fixed point but moves with a velocity p_y/m parallel to the y-axis.

To determine y and z we use

$$\dot{y} = -\frac{\partial H}{\partial p_y} = -\frac{1}{m}(p_y - qxB) = \omega(x - x_0) = a\omega\cos(\omega t + \alpha) \tag{12}$$

$$\dot{z} = \frac{p_z}{m} \tag{13}$$

whence $$y = a\sin(\omega t + \alpha) + y_0 \tag{14}$$

and $$z = \frac{p_z}{m}t + z_0 \tag{15}$$

Thus, the particle moves along a spiral of radius a with its axis that would have been parallel to **B** if $p_y = 0$. The axis passes through the point (x_0, y_0) and its distance from the yz-plane is given by $x_0 = p_y/qB$.

11.12 PHASE SPACE

While using the Lagrangian formulation for describing motion of a

7. What is a constant of motion? How can one obtain the constants of motion?
8. What are the Lagrange multipliers? How are they evaluated in actual problems?
9. Show that variation and differentiation commute. When will variation and integration commute?
10. Construct a variation whose derivative is not small at points where the variation itself is infinitesimal.

PROBLEMS

1. A particle of mass m is constrained to move on the surface of a cone having semi-vertical angle α and acted upon by gravitational force. Determine the Lagrangian and obtain Lagrange's equation for r.
2. A disc is rolling down an inclined plane without slipping. Write down the equation of constraint for this. Obtain the Lagrangian and Lagrange's equations for the disc. Using Lagrange's method of undetermined multipliers, obtain expressions for the linear and angular accelerations of the disc.
3. A particle of mass m is constrained to move on the surface of a cylinder $x^2 + y^2 = r^2$. The particle is subjected to a force directed towards the origin and the magnitude of the force is proportional to the distance of the particle from the origin. Obtain the Lagrangian and Hamiltonian for this particle. Hence, show that the motion of the particle along the z-axis is simple harmonic.
4. Discuss the motion of a one-dimensional harmonic oscillator by evaluating the Hamiltonian and then using Hamilton's equations.
5. Consider a function $y(x) = x$. Construct neighbouring paths by adding a sinusoidal term to $y(x)$. Justify that the function satisfies the extremum condition.
6. Find the dimensions of a parallelopiped of maximum volume that is circumscribed by (a) a sphere of radius r, and (b) an ellipsoid of semi-axes a, b and c.
7. A body is released from a height of 19.6 m and it strikes the ground in 2 seconds. The equation for distance h of fall during time t could hypothetically have any of the forms

$$h = gt, \quad h = \tfrac{1}{2}gt^2 \quad \text{and} \quad h = \tfrac{1}{4}gt^3$$

Show that the correct form leads to a minimum for the integral in Hamilton's Principle.
8. A particle of mass m moves under the action of gravity along a spiral $z = k\theta$, $r =$ constant, where k is constant and z is vertical. Obtain the Hamiltonian equations of motion.

9. Obtain Euler's equations for making

$$\int_{t_1}^{t_2} f(q, \dot{q}, \ddot{q}, t)\, dt$$

an extremum.

10. Obtain Euler's equations that result when Hamilton's principle is applied to a particle moving in a potential

$$V = \tfrac{1}{2}kx^2$$

11. A two-dimensional anisotropic oscillator has a potential energy $(\frac{1}{2}k_1x^2 + \frac{1}{2}k_2y^2)$. Obtain the Hamiltonian and Hamilton's equations of motion for this oscillator.

12. Write down the Hamiltonian for a spherical pendulum. Obtain Hamilton's equations therefrom.

13. Two particles of masses m_1 and m_2 move under their mutual gravitational attraction in an external gravitational field whose acceleration is g. Write down the Hamiltonian and obtain Hamilton's equations of motion.

14. Obtain the Hamiltonian of a heavy symmetrical top with one point fixed. Also obtain Hamilton's equations and solve them.

15. A particle of mass m moves in three dimensions under a conservative force with potential energy $V(r)$. Find the Hamiltonian function in terms of spherical polar coordinates. Which coordinate is ignorable?

16. For a symmetric top, express the Lagrangian in the form

$$L = \tfrac{1}{2}I_1(\dot{\theta}^2 + \dot{\varphi}^2 \sin^2 \theta) + \tfrac{1}{2}I_3(\dot{\psi} + \dot{\varphi} \cos \theta)^2 - Mgl \cos \theta$$

Obtain H. Write down Lagrange's and Hamilton's equations of motion. What are integrals of motion?

17. A heavy particle is placed at the top of a vertical hoop. Calculate the reaction of the hoop on the particle by means of Lagrange's method of undetermined multipliers and Lagrange's equations. Find the height at which the particle falls off.

18. It sometimes occurs that the generalised coordinates appear separately in the kinetic energy and the potential energy such that we have

$$T = \sum_i f_i(q_i)\dot{q}_i^2, \quad V = \sum_i V_i(q_i)$$

Show that Lagrange's equations then separate and that the problem can then be reduced to quadratures.

19. Find the brachistochrone for a field in which the potential energy of the particle decreases as the vertical distance through which the particle descends from its initial position at rest.

20. Obtain the partial differential equation describing the surface assumed by a soap film held by a wire bent in the form of a simple closed curve. The surface tension tends to minimize the area of the film.

Differentiation of equation (12.1) with respect to $\dot{q}_j$ gives

$$\frac{\partial}{\partial \dot{q}_j}\frac{df}{dt} = \frac{\partial f}{\partial q_j} \tag{12.2}$$

and further differentiation with respect to t gives

$$\frac{d}{dt}\frac{\partial}{\partial \dot{q}_j}\left(\frac{df}{dt}\right) = \frac{d}{dt}\left(\frac{\partial f}{\partial q_j}\right) = \frac{\partial}{\partial q_j}\left(\frac{df}{dt}\right) \tag{12.3}$$

where we have interchanged the order of differentiation. Equation (12.3) is Lagrange's equation satisfied by $\frac{df}{dt}$. Thus, we can define a new Lagrangian

$$L' = L + \frac{df(q_1, q_2, \ldots, q_n, t)}{dt} \tag{12.4}$$

which will always satisfy Lagrange's equations identically. Thus, the Lagrangian is not unique but is always uncertain by a term $\frac{df}{dt}$. The transformation of the Lagrangian defined in equation (12.4) which keeps the dynamical equations of motion unchanged is known as 'gauge transformation'.

In this transformation, canonical momenta are also changed. Thus, the new canonical momenta are

$$p'_j = \frac{\partial L'}{\partial \dot{q}_j} = \frac{\partial L}{\partial \dot{q}_j} + \frac{\partial}{\partial \dot{q}_j}\frac{df}{dt} = p_j + \frac{\partial f}{\partial q_j} \tag{12.5}$$

by equation (12.2).

We now introduce a coordinate transformation from old coordinates q_j to a new set of coordinates Q_k ($k = 1, 2, \ldots, n$);

$$q_j = q_j(Q_1, Q_2, \ldots, Q_n, t) \tag{12.6}$$

which is known as point transformation.

The expression $\frac{df(Q_k, t)}{dt}$ satisfies Lagrange's final equations in the new coordinates in which the Lagrangian is

$$L'(Q_k, \dot{Q}_k, t) = L(q_j, \dot{q}_j, t) \pm \frac{df(Q_k, t)}{dt} \tag{12.7}$$

In order to transform variables q_i to Q_i, we can take the arbitrary function in equation (12.7) as a function of both the new and old coordinates. Thus, $F \equiv F(q_i, Q_i, t)$ is a function of $2n$ variables besides time. Out of these $2n$ variables, however, only n variables are independent and we can write, by using the transformation equation (12.6), as

$$\begin{aligned} F = (q_j, Q_k, t) &= F[q_j(Q_1, Q_2, \ldots, Q_n, t), Q_k, t] \\ &= F(Q_1, Q_2, \ldots, Q_n, t) \end{aligned}$$

Then, transformation equation (12.7) could be written in the form

$$L'(Q_k, \dot{Q}_k, t) = L(q_j, \dot{q}_j, t) \pm \frac{dF(q_j, Q_k, t)}{dt} \tag{12.8}$$

With this substitution, equation (12.29) becomes

$$p = \sqrt{2Pm\omega}\cos Q \tag{12.32}$$

Since F_1 does not involve t explicitly, the Hamiltonian is unaffected by the transformation. In order to express H in terms of Q and P, we proceed as follows:

The Hamiltonian H for the oscillator is

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2}q^2 \tag{12.33}$$

Substituting for p and q, we get

$$H = \omega P\cos^2 Q + \omega P\sin^2 Q = \omega P \tag{12.34}$$

Equation (12.34) shows that the Hamiltonian is cyclic in Q. Hence, the conjugate momentum P must be constant.

Now $$P = \frac{H}{\omega} = \frac{E}{\omega}$$

where E is the total energy.

Further $$\dot{Q} = \frac{\partial H}{\partial P} = \omega.$$

Hence, $Q = \omega t + \alpha$, where α is a constant of integration. Thus, the transformation has changed the problem of the oscillator in such a way that the new P is a constant of motion and the new coordinate Q increases linearly with t. The equation for Q gives translational motion of the oscillator and the transformation is equivalent to changing the oscillatory motion into a translational one.

Since $$q = \sqrt{\frac{2P}{m\omega}}\sin Q$$

we get the solution

$$q = \sqrt{\frac{2E}{m\omega^2}}\sin(\omega t + \alpha) \tag{12.35}$$

This is the usual solution for a harmonic oscillator.

If generating function F_2 is chosen, then the transformation from (q, Q) to (q, P) has to be carried out. Since, we have by equation (12.27b)

$$\frac{\partial F_1}{\partial Q_k} = -P_k$$

generating function F_2 can be written as

$$F_2(q, P, t) = F_1(q, Q, t) + \sum P_k Q_k \tag{12.36}$$

Equation (12.36) can be solved for F_1. Substituting this value of F_1 in equation (12.9), wherein we write L and L' in terms of H and K, we get

$$\sum p_i \dot{q}_i - H = \sum P_i\dot{Q}_i - K + \frac{d}{dt}\{F_2(q, P, t) - \sum P_i Q_i\}$$

or $$K = H - \sum \dot{P}_i Q_i - \sum p_i\dot{q}_i + \frac{d}{dt}F_2(q, P, t) \tag{12.37}$$

Expressing the total time derivative $\frac{dF_2(q, P, t)}{dt}$ in terms of the derivatives of its argument, we obtain

$$K = H - \sum p_i \dot{q}_i - \sum \dot{P}_i Q_i + \sum \frac{\partial F_2}{\partial q_i} \dot{q}_i + \sum \frac{\partial F_2}{\partial P_i} \dot{P}_i + \frac{\partial F_2}{\partial t}$$

or
$$K = H + \frac{\partial F_2}{\partial t} - \sum \left(p_i - \frac{\partial F_2}{\partial q_i}\right) \dot{q}_i - \left(Q_i - \frac{\partial F_2}{\partial P_i}\right) \dot{P}_i \tag{12.38}$$

As we are making q and P independent variables, the coefficients of $\dot{q}_i$ and $\dot{P}_i$ must be identically zero. Thus, we get

$$p_i = +\frac{\partial F_2}{\partial q_i} \tag{12.39a}$$

$$Q_i = +\frac{\partial F_2}{\partial P_i} \tag{12.39b}$$

and
$$K = H + \frac{\partial F_2}{\partial t} \tag{12.39c}$$

By following the same procedure as in the previous two cases, we get for F_3 and F_4

$$q_i = -\frac{\partial F_3}{\partial p_i} \tag{12.40a}$$

$$P_i = -\frac{\partial F_3}{\partial Q_i} \tag{12.40b}$$

and
$$K = H + \frac{\partial F_3}{\partial t} \tag{12.40c}$$

and
$$q_i = -\frac{\partial F_4}{\partial p_i} \tag{12.41a}$$

$$Q_i = \frac{\partial F_4}{\partial p_i} \tag{12.41b}$$

and
$$K = H + \frac{\partial F_4}{\partial t} \tag{12.41c}$$

12.3 CONDITION FOR TRANSFORMATION TO BE CANONICAL

It can be shown that a transformation

$$P_l = P_l(q_k, p_k, t),\ Q_k = Q_k(q_k, p_k, t)$$

is canonical only if the expression

$$\sum_i p_i\, dq_i - \sum_i P_i\, dQ_i \tag{12.42}$$

is an exact differential.

For example, consider the generating function which transforms variables q_i, p_i to variables Q_i, P_i when time is held fixed:

$$F_1 = F_1(q_i, Q_i)$$

For this function, we have proved that

$$p_i = \frac{\partial F_1}{\partial q_i}$$

The transformation is canonical if $(p\,dq - P\,dQ)$ is an exact differential. In the present case

$$\begin{aligned} p\,dq - P\,dQ &= p\,dq - \tfrac{1}{2}(p^2+q^2)\,\frac{p\,dq - q\,dp}{p^2+q^2} \\ &= p\,dq - \tfrac{1}{2}(p\,dq - q\,dp) \\ &= \tfrac{1}{2}(p\,dq + q\,dp) \\ &= d(\tfrac{1}{2}pq) \end{aligned} \tag{12.48}$$

Thus, $p\,dq - P\,dQ$ is an exact differential and hence the transformation is canonical.

4. We now show that $\sum_k q_k Q_k$ generates the exchange transformation in which position coordinates and the momenta can be interchanged. The given generating function is of the type

$$F_1 = \sum_k q_k Q_k$$

Hence, from equation (12.27),

$$p_k = \frac{\partial F_1}{\partial q_k} = Q_k \tag{12.49}$$

and

$$P_k = -\frac{\partial F_1}{\partial Q_k} = -q_k \tag{12.50}$$

Thus, the coordinates and the momenta are interchanged by the transformation.

5. In the case of canonical transformations given by equation (12.17), we can obtain the following relations:

(i) $\dfrac{\partial q_j}{\partial Q_k} = \dfrac{\partial P_k}{\partial p_j}$, (ii) $\dfrac{\partial q_j}{\partial P_k} = -\dfrac{\partial Q_k}{\partial p_j}$

To prove relation (i), we use

$$q_j = -\frac{\partial F_3}{\partial p_j} \text{ and } P_k = -\frac{\partial F_3}{\partial Q_k}$$

from equations (12.40a) and (12.40b).

Hence

$$\frac{\partial q_j}{\partial Q_k} = -\frac{\partial^2 F_3}{\partial Q_k\,\partial p_j}$$

and

$$\frac{\partial P_k}{\partial p_j} = -\frac{\partial^2 F_3}{\partial p_j\,\partial Q_k}$$

Hence

$$\frac{\partial q_j}{\partial Q_k} = \frac{\partial P_k}{\partial p_j} \tag{12.51}$$

Relation (ii) can be proved in a similar manner using equations (12.41a) and (12.41b).

Two more relations of this type can also be proved. These are

$$\frac{\partial p_j}{\partial Q_k} = -\frac{\partial P_k}{\partial q_j} \text{ and } \frac{\partial p_j}{\partial P_k} = \frac{\partial Q_k}{\partial q_j} \tag{12.52}$$

12.5 POISSON BRACKETS

We now consider the useful representation of Poisson brackets in which the equations of motion can be written in a symmetric form. The Poisson brackets are found to be a very useful tool in quantum mechanics and field theory.

The Poisson brackets are defined by the equation

$$[u, v]_{q,p} = \sum_k \left(\frac{\partial u}{\partial q_k}\frac{\partial v}{\partial p_k} - \frac{\partial u}{\partial p_k}\frac{\partial v}{\partial q_k}\right) \tag{12.53a}$$

or

$$[u, v]_{Q,P} = \sum_k \left(\frac{\partial u}{\partial Q_k}\frac{\partial v}{\partial P_k} - \frac{\partial u}{\partial P_k}\frac{\partial v}{\partial Q_k}\right) \tag{12.53b}$$

It is obvious from the definition of the Poisson brackets that

$$[u, v] = -[v, u] \tag{12.54}$$

i.e., the Poisson brackets are anti-commutative.

Similarly, the Poisson bracket of a function with itself is identically zero. Thus

$$[u, u] = 0, \quad [v, v] = 0 \tag{12.55}$$

Moreover

$$[u, c] = 0 = [v, c] \tag{12.56}$$

where c is independent of q or p.

The Poisson bracket obeys the distributive law of algebra, viz.

$$[u + v, w] = [u, w] + [v, w] \tag{12.57}$$

Similarly

$$[u, vw] = [u, v]w + v[u, w] \tag{12.58}$$

The above properties can be proved by using the definition and the elementary properties of differentiation and are left to the reader as an exercise.

Another important property of the Poisson brackets is

$$[q_j, p_k] = \delta_{jk}$$

where δ_{jk} is the Kronecker delta.

To prove this property, we write the expansion

$$[u, v] = \sum_l \left(\frac{\partial u}{\partial q_l}\frac{\partial v}{\partial p_l} - \frac{\partial u}{\partial p_l}\frac{\partial v}{\partial q_l}\right)$$

as

$$[q_j, p_k] = \sum_l \left(\frac{\partial q_j}{\partial q_l}\frac{\partial p_k}{\partial p_l} - \frac{\partial q_j}{\partial p_l}\frac{\partial p_k}{\partial q_l}\right)$$

But, $\dfrac{\partial q_j}{\partial q_l} = \delta_{jl}$, $\dfrac{\partial p_k}{\partial p_l} = \delta_{kl}$ and $\dfrac{\partial q_j}{\partial p_l} = 0 = \dfrac{\partial p_k}{\partial q_l}$, and $\sum_l \delta_{jl}\delta_{kl} = \delta_{jk}$

Hence, we are left with

$$[q_j, p_k] = \delta_{jk} \tag{12.59}$$

It can also be proved that

$$[u, q_j] = -\frac{\partial u}{\partial p_j} \text{ and } [u, p_j] = +\frac{\partial u}{\partial q_j} \tag{12.60}$$

Let us consider a very important identity called Jacobi's identity

where we have used equation (12.60). Similarly, when we take $w = q_j$, we get, by using equation (12.60),

$$\begin{aligned} B_j &= [u, [v, q_j]] - [v, [u, q_j]] \\ &= -\left[u, \frac{\partial v}{\partial p_j}\right] + \left[v, \frac{\partial u}{\partial p_j}\right] \\ &= -\frac{\partial}{\partial p_j}[u, v] \end{aligned} \tag{12.67}$$

Substituting coefficients A_j and B_j in equation (12.65), we get

$$\begin{aligned} [u, [v, w]] + [v, [w, u]] &= \sum_j \left(\frac{\partial w}{\partial p_j}\frac{\partial [u, v]}{\partial q_j} - \frac{\partial w}{\partial q_j}\frac{\partial [u, v]}{\partial p_j}\right) \\ &= -[w, [u, v]], \end{aligned}$$

the Jacobi's identity.

We now prove that the Poisson brackets are also invariant under canonical transformations. Let F and G be any two arbitrary functions. Then,

$$[F, G]_{q,p} = \sum_j \frac{\partial F}{\partial q_j}\frac{\partial G}{\partial p_j} - \frac{\partial F}{\partial p_j}\frac{\partial G}{\partial q_j} \tag{12.68}$$

Suppose that q_j, p_j are functions of new coordinates Q_k, P_k, then equation (12.68) becomes

$$\begin{aligned} [F, G]_{q,p} &= \sum_{jk} \left[\frac{\partial F}{\partial q_j}\left(\frac{\partial G}{\partial Q_k}\frac{\partial Q_k}{\partial p_j} + \frac{\partial G}{\partial P_k}\frac{\partial P_k}{\partial p_j}\right) \right. \\ &\qquad \left. - \frac{\partial F}{\partial p_j}\left(\frac{\partial G}{\partial Q_k}\frac{\partial Q_k}{\partial q_j} + \frac{\partial G}{\partial P_k}\frac{\partial P_k}{\partial q_j}\right)\right] \\ &= \sum_k \left\{\frac{\partial G}{\partial Q_k}[F, Q_k]_{q,p} + \frac{\partial G}{\partial P_k}[F, P_k]_{q,p}\right\} \end{aligned} \tag{12.69}$$

When we consider a special case of $F = Q_k$, we get by changing G to F from equation (12.69)

$$\begin{aligned} [Q_k, F]_{q,p} &= \sum_j \frac{\partial F}{\partial Q_j}[Q_k, Q_j] + \sum_j \frac{\partial F}{\partial P_j}[Q_k, P_j] \\ &= \sum_j \frac{\partial F}{\partial P_j}\delta_{jk} \end{aligned} \tag{12.70}$$

where we have used equation (12.59)

or
$$[F, Q_k] = -\frac{\partial F}{\partial P_k} \tag{12.71}$$

Similarly, we can prove that

$$[F, P_k] = \frac{\partial F}{\partial Q_k} \tag{12.72}$$

Hence, the expression for $[F, G]_{q,p}$ becomes

$$[F, G]_{q,p} = \sum_k \left(\frac{\partial F}{\partial Q_k}\frac{\partial G}{\partial P_k} - \frac{\partial G}{\partial Q_k}\frac{\partial F}{\partial P_k}\right) = [F, G]_{Q,P} \tag{12.73}$$

Equation (12.73) shows that the Poisson bracket is also invariant under

canonical transformation in the phase space. Hence, there is no need of writing the subscripts (q, p) or (Q, P) on the Poisson brackets.

12.6 CANONICAL EQUATIONS IN TERMS OF POISSON BRACKET NOTATION

Let us now write the canonical equations in terms of the notation of the Poisson brackets. For this, we choose function F in equations (12.71) and (12.72) as the Hamiltonian H of the system. Then, we have

$$[q_i, H] = \frac{\partial H}{\partial p_i} = \dot{q}_i \tag{12.74}$$

and

$$[p_i, H] = -\frac{\partial H}{\partial q_i} = \dot{p}_i \tag{12.75}$$

The equations are now perfectly symmetric and the difference of negative sign which was present in Hamilton's equations is also eliminated.

Now, let us write the total time derivative of a function $u(q, p, t)$ as

$$\begin{aligned}\frac{du}{dt} &= \sum_i \left(\frac{\partial u}{\partial q_i}\dot{q}_i + \frac{\partial u}{\partial p_i}\dot{p}_i\right) + \frac{\partial u}{\partial t}\\ &= \sum_i \left(\frac{\partial u}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial u}{\partial p_i}\frac{\partial H}{\partial q_i}\right) + \frac{\partial u}{\partial t}\end{aligned}$$

by virtue of equations (12.74) and (12.75). Thus

$$\frac{du}{dt} = [u, H] + \frac{\partial u}{\partial t} \tag{12.76}$$

In case we take the function $u = H$, then we have

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \tag{12.77}$$

since $[H, H] = 0$.

Equation (12.76) shows that for systems in which t does not occur explicitly in the quantities under consideration, the total time derivative of such a quantity is just the Poisson bracket of the quantity with H. Thus, all the quantities having zero value for their Poisson brackets with H are, therefore, constants of the motion. Conversely, the Poisson bracket of a constant of motion with H will be zero. These properties can be used for finding out the constants of motion. From Jacobi's identity, we know that

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$$

If u and v are two constants of the motion and if w is the Hamiltonian H, we can write

$$[u, [v, H]] + [v, [H, u]] + [H, [u, v]] = 0$$

The first two terms of the identity vanish since the Poisson bracket of any constant of motion with the Hamiltonian vanishes. Thus, we are left with

$$[H, [u, v]] = 0$$

Thus, $[u, v]$ must also be a constant of motion. Thus, the Poisson bracket of two constants of the motion is itself a constant of the motion. This result is often referred to as Poisson's theorem,

12.7 INFINITESIMAL TRANSFORMATION

We now introduce the concept of infinitesimal transformations. In such a transformation, the new coordinates differ from the old coordinates by infinitesimal amounts. Hence, we would retain only first-order terms indicating changes.

Let,

$$Q_i = q_i + \delta q_i \tag{12.78}$$

and

$$P_i = p_i + \delta p_i \tag{12.79}$$

be the transformation equations wherein δq_i and δp_i are the infinitesimal changes in coordinates and momenta. It is obvious that the generating function will also differ only by an infinitesimal amount from the one corresponding to an identity transformation given by equation (12.45). Thus, we can write

$$F_2 = \sum_i q_i P_i + \epsilon G(q, p) \tag{12.80}$$

where ϵ is some infinitesimal parameter of transformation. Then, we can write by equations (12.39a)

$$\frac{\partial F_2}{\partial q_i} = p_i = P_i + \epsilon \frac{\partial G}{\partial q_i}$$

or

$$P_i - p_i = \delta p_i = -\epsilon \frac{\partial G}{\partial q_i} \tag{12.81}$$

Similarly, by equation (12.39b), we have

$$Q_i = \frac{\partial F_2}{\partial P_i} = q_i + \epsilon \frac{\partial G}{\partial P_i} \tag{12.82}$$

The term $\epsilon \dfrac{\partial G}{\partial P_i}$ can be replaced by a nearly equal term, viz. $\epsilon \dfrac{\partial G}{\partial p_i}$ since $P_i - p_i$ is very small and further the term under consideration is of the first order in ϵ. The difference involved in this replacement is of the second order of smallness.

Thus, equation (12.82) becomes

$$\delta q_i = Q_i - q_i = \epsilon \frac{\partial G}{\partial p_i} \tag{12.83}$$

As an illustration, consider an infinitesimal canonical transformation in which G is the Hamiltonian $H(q, p)$ and ϵ is an infinitesimal interval of time dt. Then, the corresponding changes in the coordinates and the momenta are given by

$$\delta q_i = dt \frac{\partial H}{\partial p_i} = dt\, \dot{q}_i = dq_i \tag{12.84}$$

and

$$\delta p_i = -dt \frac{\partial H}{\partial q_i} = dt\, \dot{p}_i = dp_i \tag{12.85}$$

that coordinate q_l is cyclic. Then, the Hamiltonian will be independent of q_l. It will, moreover, be invariant under an infinitesimal contact transformation which involves a displacement in q_l alone. Then, the transformation equations would take the form

$$\left.\begin{aligned} \delta q_j &= \epsilon\delta_{lj} \\ \text{and} \qquad \delta p_j &= 0 \end{aligned}\right\} \tag{12.88}$$

Thus, ϵ is an infinitesimal displacement of q_l. The generating function which produces such a transformation and satisfying equation (12.88) is

$$G = p_l \tag{12.89}$$

Hence, generating function G must be a constant of the motion, since the infinitesimal canonical transformation leaves the Hamiltonian unchanged.

Thus, p_l—the momentum conjugate to q_l—is a constant of the motion. This is the conservation of linear momentum.

Suppose that an infinitesimal contact transformation of the dynamical variables produces rotation $d\theta$ of the system. This could be imagined in yet another way. Here, we assume that the system is held stationary and the coordinate axes are rotated through angle $(-d\theta)$. If the rotation takes place, say, about the z-axis, the new coordinates are

$$X_i = x_i - y_i\, d\theta \tag{12.90}$$

$$Y_i = q_i + x_i\, d\theta \tag{12.91}$$

and

$$Z_i = z_i \tag{12.92}$$

Hence, the infinitesimal changes in the coordinates are

$$\delta x_i = -y_i\, d\theta, \quad \delta y_i = x_i\, d\theta, \quad \delta z_i = 0 \tag{12.93}$$

Similar equations involving the changes in the components of momentum are

$$\delta p_{ix} = -p_{iy}\, d\theta, \quad \delta p_{iy} = p_{ix}\, d\theta, \quad \delta p_{iz} = 0. \tag{12.94}$$

The generating function which will yield equations (12.93) with the use of equations (12.81) and (12.83) is

$$G = \sum_i (x_i p_{iy} - y_i p_{ix}) \tag{12.95}$$

The value of the parameter ϵ in the present case will be $d\theta$. We can compute the values of δx_i, δy_i, δp_{ix} and δp_{iy} from generating function G by using equations (12.81) and (12.83). Since the rotation is about the z-axis, we can also write

$$G = L_z$$

This last result can be generalised and written as

$$G = \mathbf{L}\cdot\hat{\mathbf{e}} \tag{12.96}$$

where $\hat{\mathbf{e}}$ is a unit vector along the direction of the infinitesimal rotation vector. Thus, angular momentum is the generator of the infinitesimal rotational motion of the system.

By using the properties of the Poisson brackets we can verify the relation

$$[L_x, L_y] = L_z \tag{12.97}$$

or in a general notation as

$$[L_i, L_j] = \mathcal{E}_{ijk} L_k \tag{12.98}$$

where $i, j, k = 1, 2, 3$.

If L_x and L_y are constants of motion, equation (12.97) states that L_z is also a constant of the motion. Thus, if any two components of the angular momentum are constant, the total angular momentum vector is conserved.

Similarly, the relation

$$[L^2, L_i] = 0, \quad \text{for } i = 1, 2, 3 \tag{12.99}$$

can be proved in a straightforward way.

We have already remarked that $[p_i, p_j] = 0$. But, equation (12.98) shows that $[L_i, L_j] \neq 0$. Hence, if one component of the angular momentum along a fixed direction is taken as a canonical momentum, the two perpendicular components cannot be simultaneously the canonical momenta. But, equation (12.99) states that the scalar magnitude of the angular momentum and anyone of its components can be simultaneously canonical.

12.9 THE HAMILTON-JACOBI EQUATIONS

We have seen that if the Hamiltonian H is conserved, we can solve the problem by transforming to new canonical coordinates which would be all cyclic. The equations in the new form are then easily integrable. Another way of solving the problem would be to make a canonical transformation from coordinates and momenta, (q, p) at time t to their initial values (q_0, p_0) which are known. Thus, relations of the type

$$\begin{aligned} q &= q(q_0, p_0, t) \\ \text{and} \qquad p &= p(q_0, p_0, t) \end{aligned} \tag{12.100}$$

can be obtained as solutions of the problems. This is a more general procedure and is applicable, at least in principle, to a case when the Hamiltonian involves time.

The equations of motion in terms of the transformed Hamiltonian and the relation between the new and the old Hamiltonian are

$$\begin{aligned} \frac{\partial K}{\partial P_i} &= \dot{Q}_i \\ -\frac{\partial K}{\partial Q_i} &= \dot{P}_i \\ K &= H + \frac{\partial F}{\partial t} \end{aligned} \tag{12.101}$$

where F is the generating function of the transformation. If new variables Q_i and P_i are constants of motion, i.e. $\dot{Q}_i = 0$ and $\dot{P}_i = 0$, then

$$\frac{\partial K}{\partial Q_i} = 0 \text{ and } \frac{\partial K}{\partial P_i} = 0 \tag{12.102}$$

Moreover, the new Hamiltonian K can be taken equal to zero.

Generating function F then satisfies the equation

$$H(q, p, t) + \frac{\partial F}{\partial t} = 0 \tag{12.103}$$

We choose the generating function to be of the form

$$F_2 = F_2(q, p, t) \tag{12.104}$$

Now, $p_i = \dfrac{\partial F_2}{\partial q_i}$ according to equation (12.39a). Hence, equation (12.103) assumes the form

$$H\left(q_1, q_2, \ldots, q_n; \frac{\partial F_2}{\partial q_1}, \frac{\partial F_2}{\partial q_2}, \ldots, \frac{\partial F_2}{\partial q_n}, t\right) + \frac{\partial F_2}{\partial t} = 0 \tag{12.105}$$

Equations (12.105) are known as the Hamilton-Jacobi equations. They are first-order partial differential equations in $(n + 1)$ variables, viz. $q_1, q_2, \ldots, q_n, t$. By convention, the solution of equation (12.105) is denoted by S and is called Hamilton's principal function.

In the subsequent description, we shall use the symbol S instead of F_2.

Since equation (12.105) is a first-order partial differential equation in $(n + 1)$ variables, its complete solution must involve $(n + 1)$ constants of integration say $\alpha_1, \alpha_2, \ldots, \alpha_{n+1}$. However, S does not appear directly in equation (12.105); but appears as partial derivatives with respect to q or t. Further, if S is a solution of equation (12.105), $(S + \alpha)$ must also be a solution. Thus, out of $(n + 1)$ constants mentioned above, one is an additive constant attached to S. The complete solution of equation (12.105) is, therefore, of the form

$$S = S(q_1, q_2, \ldots, q_n, \alpha_1, \alpha_2, \ldots, \alpha_n, t) \tag{12.106}$$

where none of the n constants $\alpha_1, \ldots, \alpha_n$ is additive. Thus, S is a function of n coordinates q_i, n constants α_i, and time. This is precisely the same description as that of the generating function. Constants α_i can be chosen as new momenta P_i which are known to be constants. Thus

$$P_i = \alpha_i,\ i = 1, 2, \ldots, n \tag{12.107}$$

and the new momenta are chosen to be the values of p_0 at initial time t_0. We can now write the transformation equations as

$$p_i = \frac{\partial S}{\partial q_i} \tag{12.108}$$

Equation (12.108) yields relations between α_i and p_i and q_i at time t_0.

Further, the transformation equations

$$Q_i = \frac{\partial S}{\partial P_i} = \frac{\partial S}{\partial \alpha_i} = \beta_i, \text{ say} \tag{12.109}$$

where n values of β_i may be chosen as in the case of momenta and are also expressible in terms of initial conditions. Equation (12.109) can also be written as

$$q_j = q_j(\alpha_i, \beta_i, t) \tag{12.110}$$

This along with the other relations

$$p_j = p_j(\alpha_i, \beta_i, t) \tag{12.111}$$

gives the solutions of the problem.

In order to understand the meaning of S, let us write the total time derivative of S as

$$\frac{dS}{dt} = \sum_i \frac{\partial S}{\partial q_i}\dot{q}_i + \frac{\partial S}{\partial t}$$
$$= \sum_i p_i\dot{q}_i - H = L \tag{12.112}$$

by virtue of equations (12.108) and (12.105).
Hence, Hamilton's principal function S is

$$S = \int L\,dt + \text{const.} \tag{12.113}$$

Thus, Hamilton's principal function differs from the indefinite time integral of the Lagrangian by a constant term.

As an illustration, consider the one-dimensional harmonic oscillator with the Hamiltonian H

$$H = \frac{p^2}{2m} + \frac{m\omega^2 q^2}{2}, \quad \omega = \sqrt{\frac{k}{m}} \tag{12.114}$$

which is independent of time. Now, $p = \dfrac{\partial S}{\partial q}$. Hence

$$H = \frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{m\omega^2 q^2}{2} \tag{12.115}$$

With this expression for H, the Hamilton-Jacobi equation (12.105) can be written as

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{m\omega^2 q^2}{2} + \frac{\partial S}{\partial t} = 0 \tag{12.116}$$

Since, the only term that involves an explicit dependence of S on t is the last term, the solution of equation (12.116) can be found out in the form

$$S(q, \alpha, t) = W(q, \alpha) - \alpha t \tag{12.117}$$

where α is a constant of integration. Then, equation (12.116) takes up the form

$$\frac{1}{2m}\left(\frac{\partial W}{\partial q}\right)^2 + \frac{m\omega^2 q^2}{2} = \alpha \tag{12.118}$$

or
$$\frac{\partial W}{\partial q} = \sqrt{2m\left(\alpha - \frac{m\omega^2 q^2}{2}\right)} = m\omega\sqrt{\left(\frac{2\alpha}{m\omega^2} - q^2\right)}$$

On integration, this yields

$$W = m\omega \int \sqrt{\left(\frac{2\alpha}{m\omega^2} - q^2\right)}\,dq \tag{12.119}$$

Expressing in terms of S, we get

$$S = W - \alpha t$$
$$= m\omega \int \sqrt{\left(\frac{2\alpha}{m\omega^2} - q^2\right)}\,dq - \alpha t \tag{12.120}$$

We, however, do not need the value of S. What we need is $\dfrac{\partial S}{\partial \alpha}$. Now,

the solution for q arises out of equation (12.109). Thus

$$\beta = \frac{\partial S}{\partial \alpha} = \frac{1}{\omega}\int \frac{dq}{\sqrt{\left(\frac{2\alpha}{m\omega^2} - q^2\right)}} - t$$

Hence
$$t + \beta = -\frac{1}{\omega} \text{ arc cos } q\sqrt{\frac{m\omega^2}{2\alpha}}$$

Then, we can write

$$q = \sqrt{\frac{2\alpha}{m\omega^2}} \cos \omega(t + \beta) \qquad (12.121)$$

This is the usual form of the solution for a one-dimensional harmonic oscillator.

Constants α and β will now be related to initial values q_0 and p_0. Let the particle be at rest at $t = 0$, so that $p_0 = 0$; but it is displaced from its equilibrium position by amount q_0. From equation (12.116), substituting and simplifying, we get

$$\left(\frac{\partial S}{\partial q}\right)_0 = p_0 = \sqrt{2m}\left(\alpha - \frac{m\omega^2 q_0^2}{2}\right)^{1/2}$$

But $p_0 = 0$.

Hence, we get
$$\alpha = \frac{m\omega^2 q_0^2}{2} \qquad (12.122)$$

which is, according to equation (12.114), the initial total energy of the system. From equation (12.122), $\sqrt{\frac{2\alpha}{m\omega^2}} = q_0$. Hence, equation (12.121) becomes

$$q = q_0 \cos \omega(t + \beta) \qquad (12.123)$$

Since, at $t = 0$, $q = q_0$, we observe that $\beta = 0$.

Thus, Hamilton's principal function S is the generator of a contact transformation transforming harmonic oscillator with a canonical momentum $\alpha = H$, the total energy, and a coordinate β which vanishes initially at $t = 0$.

Hamilton's principal function can be written as

$$S = m\omega \int \sqrt{q_0^2 - q^2}\, dq - \frac{m\omega^2 q_0^2 t}{2}$$

by substituting for α from equation (12.122) in equation (12.120).

Or
$$S = m\omega q_0^2 \int (\sin^2 \omega t - \tfrac{1}{2})\, dt$$

by putting q from equation (12.123).

It should be noted that in writing this last step, we have taken the negative square root, viz.

$$\sqrt{q_0^2 - q^2} = -\sin \omega t$$

Now, the Lagrangian is

$$L = \tfrac{1}{2}m\dot{q}^2 - \frac{m\omega^2 q^2}{2}$$

$$= \frac{m\omega^2 q_0^2}{2}(\sin^2 \omega t - \cos^2 \omega t)$$

$$= m\omega^2 q_0^2 (\sin^2 \omega t - \tfrac{1}{2}) \tag{12.124}$$

Thus, S is the time integral of the Lagrangian.

A particular case of interest is when function S could be separated into two parts—one involving coordinate q only and the other involving time t only. This kind of separation of variables and hence the integration of the Hamilton-Jacobi equation is possible whenever the old Hamiltonian does not involve the time explicitly. The Hamilton-Jacobi equation can then be written down as

$$H\left(q_i, \frac{\partial S}{\partial q_i}\right) + \frac{\partial S}{\partial t} = 0 \tag{12.125}$$

It will be observed that the first term involves the dependence of S on q while the second term depends on t.

The solution S will be of the type

$$S(q_i, \alpha_i, t) = W(q_i, \alpha_i) - \alpha_1 t \tag{12.126}$$

With this, differential equation (12.125) can be written as

$$H\left(q_i, \frac{\partial W}{\partial q_i}\right) = \alpha_1 = \text{const.} \tag{12.127}$$

which does not involve time at all. Constant α_1, which also appears as one of the constants in S is, by equation (12.127), equal to constant value H.

Although, time independent function W appears as a part of S, it can be shown that W separately generates its own contact transformation that has properties very much different from that generated by S.

Consider a canonical transformation in which the new momenta are all constants of the motion α_i, and α_1 is equal to H. Let $W(q, p)$ be the generating function for this transformation. Then, the transformation equations are

$$p_i = \frac{\partial W}{\partial q_i} \quad \text{and} \quad Q_i = \frac{\partial W}{\partial P_i} = \frac{\partial W}{\partial \alpha_i} \tag{12.128}$$

The condition that

$$H(q_i, p_i) = \alpha_1 \tag{12.129}$$

must hold in addition to equation (12.128).

With the substitution of $p_i = \dfrac{\partial W}{\partial q_i}$, equation (12.129) becomes

$$H\left(q_i, \frac{\partial W}{\partial q_i}\right) = \alpha_1 \tag{12.130}$$

It is observed that equation (12.117) is identical with equation (12.127).

But, W does not involve time. Hence, the old and new Hamiltonians are equal. Thus

$$K = \alpha_1 \tag{12.131}$$

Function W is known as Hamilton's characteristic function. It generates a canonical transformation in which all the new coordinates are cyclic. The canonical equations for P_i and Q_i are

$$\dot{P}_i = -\frac{\partial K}{\partial Q_i} \quad \text{or} \quad P_i = \alpha_i \tag{12.132}$$

and

$$\left.\begin{aligned} \dot{Q}_i = \frac{\partial K}{\partial P_i} = \frac{\partial K}{\partial \alpha_i} &= 1 \quad \text{for } i = 1 \\ &= 0 \quad \text{for } i \neq 1 \end{aligned}\right\} \tag{12.133}$$

The solutions of equation (12.133) are then

$$\begin{aligned} Q_1 &= t + \beta_1 \\ &= \frac{\partial W}{\partial \alpha_1}, \text{ by equation (12.128)} \end{aligned}$$

and

$$\begin{aligned} Q_i &= \beta_i \\ &= \frac{\partial W}{\partial \alpha_i}, \ i \neq 1, \text{ by equation (12.128)} \end{aligned}$$

Thus, Q_1 is the only coordinate which is not a constant of motion. Q_1 is related with time and we have here an example of time t and Hamiltonian H as canonically conjugate variables.

The total time derivative of Hamilton's characteristic function $W = W(q_i, \alpha_i)$ gives

$$\frac{dW}{dt} = \sum \frac{\partial W}{\partial q_i} \dot{q}_i = \sum p_i \dot{q}_i$$

and hence

$$W = \int \sum p_i \dot{q}_i \, dt = \int \sum p_i \, dq_i$$

which is called the action of the system. In the case of a simple harmonic oscillator, the action $\int p \, dq$ is the area of the ellipse $2\pi E/\omega$ and is a constant quantity. Classically, the oscillator can be given oscillations with any amplitude and hence E can change continuously.

In quantum mechanics, however, the action (energy) of an oscillator is a quantised quantity and hence the theory given above forms a useful way of departure from classical to quantum mechanical problems.

12.10 SEPARATION OF VARIABLES

We can separate the variables in the Hamilton-Jacobi equation under suitable conditions. Whenever this is possible, the Hamilton-Jacobi method becomes extremely useful.

Consider systems in which the Hamiltonian is one of the constants of the motion. It should be noted that it may not necessarily be the total energy. In such a case, we need consider only the contact transformation

generated by Hamilton's characteristic function W and its corresponding Hamilton-Jacobi equation.

Variables q_i occurring in the equation are said to be separable if the solution of the form

$$W = \sum_i (q_i, \alpha_1, \alpha_2, \ldots, \alpha_n) \tag{12.134}$$

splits the Hamilton-Jacobi equation into n equations

$$H\left(q_i, \frac{\partial W_i}{\partial q_i}, \alpha_1, \alpha_2, \ldots, \alpha_n\right) = \alpha_i \tag{12.135}$$

Only one of the coordinates q_i and partial derivative of W_i with respect to α_i is involved in each of the equations. These equations are first-order differential equations and can be solved for $\frac{\partial W_i}{\partial q_i}$ and then integrated with respect to q_i.

We have seen earlier that if H is not an explicit function of t, separation of variables is possible. The solution for S was obtained in the form

$$S(q_i, \alpha_i, t) = W(q_i, \alpha_i) + S_2(t, \alpha_i) \tag{12.136}$$

With this, the Hamilton-Jacobi equation becomes

$$H\left(q_i, \frac{\partial W}{\partial q_i}\right) + \frac{\partial S_2}{\partial t} = 0 \tag{12.137}$$

Equation (12.137) holds only if

$$\frac{\partial S_2}{\partial t} = -\alpha_1 \tag{12.138}$$

and

$$H\left(q_i, \frac{\partial W}{\partial q_i}\right) = \alpha_1 \tag{12.139}$$

Equation (12.138) gives $S_2 = -\alpha_1 t$ as obtained earlier. Equation (12.139) represents the Hamilton-Jacobi equation for function W.

A similar separation of variables is possible when all except one coordinates are cyclic. Let q_1 be a noncyclic coordinate and the solution for W be of the form

$$W = \sum_i W_i(q_i, P_i) \tag{12.140}$$

Since, all coordinates except q_1 are cyclic, we must have

$$\frac{\partial W_i}{\partial q_i} = p_i = \alpha_i, \; i \neq 1 \tag{12.141}$$

Then, the Hamilton-Jacobi equation reduces to

$$H\left(q_1, \frac{\partial W}{\partial q_1}, \alpha_2, \ldots, \alpha_n\right) = \alpha_1 \tag{12.142}$$

Equation (12.142) is an ordinary first-order differential equation for W_1 and hence it can be solved immediately. Equations (12.141) and (12.142) together completely specify function W. Integrating equation (12.141), we get

$$W_i = \alpha_i q_i, \; i \neq 1$$

Hence, we can write

$$W = W_1 + \sum_{i=2}^{n} \alpha_i q_i \tag{12.143}$$

where W_1 corresponds to $i = 1$. Equation (12.143) has the same form as that of equation (12.117).

As an illustration, consider a particle in a central force-field. The Hamiltonian H has the form

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\varphi^2}{r^2}\right) + V(r) \tag{12.144}$$

It is observed that the Hamiltonian is cyclic in φ. Hence, Hamilton's characteristic function can be written as

$$W = W_1(r) + \alpha_\varphi \varphi \tag{12.145}$$

where constant α_φ is the constant angular momentum p_φ conjugate to φ.

The Hamilton-Jacobi equation can then be written as

$$\frac{1}{2m}\left[\left(\frac{\partial W_1}{\partial r}\right)^2 + \frac{\alpha_\varphi^2}{r^2}\right] + V(r) = \alpha_1 \tag{12.146}$$

where α_1 represents the total constant energy of the system. Equation (12.146) yields

$$\frac{\partial W_1}{\partial r} = \sqrt{2m(\alpha_1 - V) - \frac{\alpha_\varphi^2}{r^2}}$$

Hence

$$W = W_1 + \alpha_\varphi \varphi = \int \sqrt{2m(\alpha_1 - V) - \frac{\alpha_\varphi^2}{r^2}}\, dr + \alpha_\varphi \varphi$$

Using this expression for Hamilton's characteristic function the transformation equations become

$$t + \beta_1 = \frac{\partial W}{\partial \alpha_1} = \int \frac{m\, dr}{\sqrt{2m(\alpha_1 - V) - \frac{\alpha_\varphi^2}{r^2}}} \tag{12.147}$$

and

$$\beta_2 = \frac{\partial W}{\partial \alpha_\varphi} = -\int \frac{\alpha_\varphi\, dr}{r^2\sqrt{2m(\alpha_1 - V) - \frac{\alpha_\varphi^2}{r^2}}} + \varphi \tag{12.148}$$

From equation (12.148), we have

$$\varphi = \beta_2 + \int \frac{\alpha_\varphi\, dr}{r^2\sqrt{2m(\alpha_1 - V) - \frac{\alpha_\varphi^2}{r^2}}}$$

Substituting $u = \frac{1}{r}$, this becomes

$$\varphi = \beta_2 - \int \frac{du}{\sqrt{\frac{2m}{\alpha_\varphi^2}(\alpha_1 - V) - u^2}}$$

This provides the orbit equation.

(ii) $q_2 = -\left(\frac{Q_1}{\omega_1}\right)^{1/2} \cos P_1 + \left(\frac{Q_2}{\omega_2}\right)^{1/2} \cos P_2$

(iii) $p_1 = (\omega_1 Q_1)^{1/2} \sin P_1 + (\omega_2 Q_2)^{1/2} \sin P_2$

(iv) $p_2 = -(\omega_1 Q_1)^{1/2} \sin P_1 + (\omega_2 Q_2)^{1/2} \sin P_2$

is a canonical transformation. If the Hamiltonian of the system is

$$H = \tfrac{1}{2}p_1^2 + \tfrac{1}{2}p_2^2 + \tfrac{1}{4}\omega_1^2(q_1 - q_2)^2 + \tfrac{1}{4}\omega^2(q_1 + q_2)^2$$

derive the new Hamiltonian $K(Q_1, Q_2, P_1, P_2)$ and solve the new Hamiltonian equations of motion. What physical system is described by this formulation?

17. A projectile is fired with velocity v_0 at angle θ with the horizontal. Assuming that the gravitational field is uniform, use the Hamilton-Jacobi method and find the equation of the trajectory and the motion of the projectile as a function of time.
18. Consider the motion of an elastic plane pendulum. An elastic pendulum can be considered as a bob attached to the end of an elastic spring with unstretched length L. Assume small oscillations and determine the frequencies of oscillation by the Hamilton-Jacobi method.
19. A homogeneous bar is free to slide on a smooth vertical plane which is constrained to rotate with angular velocity ω about a vertical axis fixed in a plane. Describe the motion by using the Hamilton-Jacobi method.
20. Prove that the Poisson bracket of two constants of motion is itself a constant of motion even when the constants depend upon time explicitly.
21. Show that if the Hamiltonian H and quantity F are constants of the motion, $\partial F/\partial t$ must also be a constant.
22. A free particle of mass m is in uniform motion. The Hamiltonian H and function $F = x - \dfrac{pt}{m}$ are constants of the motion. Show by direct computation that the constant of the motion $\partial F/\partial t$ agrees with $[H, F]$.
23. Set up the problem of a spherical pendulum in the Hamiltonian formulation using spherical polar coordinates. In terms of these canonical variables, evaluate directly the Poisson brackets
$$[L_x, L_y],\ [L_y, L_z],\ [L_z, L_x].$$
24. Discuss the problem of one-dimensional harmonic oscillator by the Hamilton-Jacobi method.
25. Set up the problem of a heavy symmetrical top, with one point fixed in the Hamilton-Jacobi method and obtain the formal solution to the motion of the top.

13
Theory of Small Oscillations

In this chapter, we shall discuss, in brief, the theory of oscillations of such small amplitudes that only the fundamental frequencies are excited. We shall limit our discussion to the theory of small oscillations about the position of stable equilibrium although it can be applied to small oscillations about the stable motion as well. We also wish to extend our discussion to systems which possess many degrees of freedom.

It is well known that the displacement $x(t)$ at any instant of a single particle, elastically coupled to a fixed support, is given by

$$x(t) = A \cos \omega t \tag{13.1}$$

where A is the amplitude and ω is the angular frequency of the motion. If a system consists of many particles coupled together, the displacement of a given particle at any instant will, in general, be a complicated function depending upon the behaviour of all the other particles to which the particle under consideration is coupled. The problem may be simplified by the transformation of the rectangular coordinates to a set of generalised coordinates q_k. It is always possible to find such a set of generalised coordinates, each of which undergoes periodic changes with a single, well-defined frequency. Then

$$q_k(t) = A \cos \omega_k t \tag{13.2}$$

Such coordinates are called the normal coordinates of the system and were introduced in the chapter on oscillations. The normal coordinates need not necessarily be the actual particle coordinates. In fact, they are, in general, complicated functions of the actual particle coordinates. The initial conditions, viz. displacement $q_k(0)$ and velocity $\dot{q}_k(0)$ at $t = 0$ can be arranged such that the subsequent motion takes place with a single frequency ω_k. Then, one of the normal modes of oscillation is said to be excited. The general motion of the system will be a complicated combination of all the normal modes.

The theory of small oscillations finds applications in a variety of fields

like acoustics, molecular spectra, coupled circuits etc. We, therefore, develop a theory of small oscillations in this chapter based on the Lagrangian formulations.

13.1 GENERAL CASE OF COUPLED OSCILLATIONS

We wish to obtain equations of motion of a system having many degrees of freedom near its equilibrium configuration. For simplicity, we consider conservative systems only. In that case, the potential energy is a function of position only. We assume that the transformation relations used to define generalised coordinates q_k of the system do not involve time explicitly. Thus, we exclude the time-dependent constraints from our discussion.

A system is said to be in equilibrium if generalised forces Q_k acting on it vanish. Since the system is conservative, forces Q_k are derivable from a potential energy function V. Hence, when the system is in equilibrium

$$Q_k = \left(\frac{\partial V}{\partial q_k}\right)_0 = 0 \tag{13.3}$$

The suffix zero is used to indicate the equilibrium configuration. Equation (13.3) shows that potential energy V has an extremum value in the equilibrium configuration of the system. We would represent the equilibrium configuration of the system by generalised coordinates q_{0k}. If, initially, the configuration is at the equilibrium position and if initial velocities $\dot{q}_{0k}$ are zero, the system would continue to remain in equilibrium indefinitely. For example, a pendulum at rest, an egg standing on its tip, etc.

The equilibrium is said to be stable if the system performs a bound motion about its equilibrium position when disturbed slightly from it. On the other hand, if the result of the small disturbance given to the system in the equilibrium configuration is the unbounded motion, the equilibrium is an unstable one. The pendulum constitutes the illustration of the stable equilibrium, while the egg standing on its tip that of the unstable equilibrium. The equilibrium is stable if the extremum value of V is a minimum, and unstable if it is a maximum.

Let us assume that during the motion of the system, the departures from the configuration of the stable equilibrium are negligibly small. We can, therefore, expand all the functions in a Taylor series about the equilibrium state and retain only the lowest order non-vanishing terms.

The potential energy function V can be expanded as

$$V(q_1, q_2, \ldots, q_n) = V(q_{01}, q_{02}, \ldots, q_{0n}) + \sum_k \left(\frac{\partial V}{\partial q_k}\right)_0 q_k + \tfrac{1}{2} \sum_{jk} \left(\frac{\partial^2 V}{\partial q_j \, \partial q_k}\right)_0 q_j q_k + \ldots \tag{13.4}$$

It should be noted that the expansion in equation (13.4) has this particular form only if $q_{0k} = 0$ and hence q_k represents displacements from

$q_{0k} = 0$. Now, $\left(\dfrac{\partial V}{\partial q_k}\right)_0 = 0$ according to equation (13.3). Further, $V(q_{01}, q_{02}, \ldots)$ is the potential energy in the equilibrium position. Without loss of generality, we can conveniently set this term equal to zero. Then, equation (13.4) reduces to

$$V = \tfrac{1}{2} \sum_{jk} \left(\frac{\partial^2 V}{\partial q_j\, \partial q_k}\right)_0 q_j q_k = \tfrac{1}{2} \sum_{jk} V_{jk} q_j q_k \tag{13.5}$$

where $V_{jk} = \left(\dfrac{\partial^2 V}{\partial q_j\, \partial q_k}\right)_0$ represent constants that depend upon the equilibrium values of q_k's, that is, on q_{0k}'s. It is obvious from the definition of V_{jk} that these constants are symmetric, i.e., $V_{jk} = V_{kj}$. In matrix notation, equation (13.5) can be written in compact form as

$$V = \tfrac{1}{2}\mathbf{q}^\dagger \mathbf{V} \mathbf{q}$$

where $\mathbf{q}$ is a column matrix and $\mathbf{q}^\dagger$ is its Hermitian adjoint.
Thus

$$\mathbf{q} = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{pmatrix} \quad \text{and} \quad \mathbf{V} = \begin{pmatrix} V_{11} & V_{12} & \ldots & V_{1n} \\ V_{21} & V_{22} & \ldots & V_{2n} \\ \vdots & \vdots & & \vdots \\ V_{n1} & V_{n2} & \ldots & V_{nn} \end{pmatrix}$$

The kinetic energy of the system will be quadratic function of velocities and by analogy with equation (8.40) can be written as

$$T = \tfrac{1}{2} \sum_{jk} m_{jk} \dot{q}_j \dot{q}_k \tag{13.6}$$

The coefficients m_{jk} are, in general, functions of coordinates q_k. These may also be expanded in a Taylor series as

$$m_{jk}(q_1, q_2, \ldots, q_n) = m_{jk}(q_{01}, q_{02}, \ldots, q_{0n}) + \sum_l \left(\frac{\partial m_{jk}}{\partial q_l}\right)_0 q_l + \ldots \tag{13.7}$$

Equation (13.6) is quadratic in $\dot{q}_k$'s. Hence, the lowest non-vanishing approximation to T is obtained by dropping all the terms except the first in the expansion of m_{jk}. Let the constant values of m_{jk}, viz. $m_{jk}(q_{01}, q_{02}, \ldots, q_{0n})$, be denoted by T_{jk}.

Hence, total kinetic energy T can be written as

$$T = \tfrac{1}{2} \sum_{jk} T_{jk} \dot{q}_j \dot{q}_k = \tfrac{1}{2}\mathbf{q}^\dagger \mathbf{V} \mathbf{q} \tag{13.8}$$

The constants T_{jk} must also be symmetric in view of their definition.

The Lagrangian L of the system can now be written as

$$\begin{aligned} L &= T - V \\ &= \tfrac{1}{2} \sum_{jk} (T_{jk} \dot{q}_j \dot{q}_k - V_{jk} q_j q_k) \end{aligned} \tag{13.9}$$

Considering q's as the generalised coordinates and using the expression for L as given in equation (13.9), Lagrange's equation yields n equations

$$\sum_j T_{jk} \ddot{q}_j + \sum_j V_{jk} q_j = 0 \tag{13.10}$$

If these n equations are solved, we will be able to describe the motion of the system near its equilibrium.

In the above equation, T_{jk} and V_{jk} are $n \times n$ arrays of numbers which specify the way in which the motions of the various coordinates are coupled. For example, if $m_{jk} \neq 0$ for $j \neq k$, then the kinetic energy would contain a term that is proportional to $\dot{q}_j \dot{q}_k$. Thus, there exists a coupling between jth and kth coordinates. If, however, m_{jk} is diagonal, so that $m_{jk} = 0$ when $j \neq k$, then the kinetic energy will be given by

$$T = \tfrac{1}{2} \sum_r m_r \dot{q}_r^2$$

where $m_r = m_{rr}$. This shows that the total kinetic energy is equal to the sum of the kinetic energies associated with various individual coordinates. Similarly, if V_{jk} is diagonal, total potential energy V is also equal to the sum of individual potential energies. Then, each coordiuate will behave in a simple manner, undergoing oscillations with a single, well-defined frequency. Hence, the problem reduces to finding out a coordinate transformation which simultaneously diagonalises T_{jk} and V_{jk}. Then, the system can be described in an extremely simple manner. The coordinates so found are called normal coordinates.

13.2 EIGENVECTORS AND EIGENFREQUENCIES

The motion we are dealing with is an oscillatory motion. Hence, solution of equation (13.10) can be expected to be of the form

$$q_j(t) = a_j \exp\left[-i(\omega t - \delta_j)\right] \tag{13.11}$$

Here, a_j are the real amplitudes of oscillation corresponding to coordinates q_j. Phase δ_j is included so as to give two constants, viz. a_j and δ_j since each of the n equations is a second order differential equation. It should be noted that the actual motion is represented by the real part of right-hand side of equation (13.11). Frequency ω and phase factor δ are determined by the equations of motion. If ω is a real quantity, then only equation (13.11) represents an oscillatory motion. The requirement that ω must be a real quantity can be argued as follows: If suppose ω contains an imaginary part, then the expression for q_j contains terms of the form $e^{\omega t}$ and $e^{-\omega t}$ Thus, when the total energy is computed, it will contain factors that increase or decrease monotonically with time. But, this is in contradiction with our assumption namely that we are dealing with a conservative system. Hence, ω must be real.

Substituting the trial solution expressed in equation (13.11) in equation (13.10), we obtain

$$\sum_j (V_{jk} - \omega^2 T_{jk}) a_j = 0 \tag{13.12}$$

Equation (13.12) represents a set of n linear homogeneous algebraic equations which must be satisfied by a_j. For nontrivial solutions to exist, the determinant of the coefficients of a_j must vanish.

Thus,

$$|V_{jk} - \omega^2 T_{jk}| = 0 \tag{13.13}$$

$$\sum_{jk} T_{jk} a_{jr} a_{ks} = \delta_{rs} \tag{13.27}$$

where δ_{rs} is the Kronecker delta.

The eigenvectors $\mathbf{a}_r$ *so defined are said to form an orthonormal set.* They are orthogonal in the sense of the condition of equation (13.23) and they are normalised in accordance with the condition mentioned in equation (13.26).

It should be noted that T_{jk} and V_{jk} are the tensor elements and equation (13.12) could have been written as

$$\mathbf{Va} = \omega^2 \mathbf{Ta}$$

The equations get quite a compact look in this notation. The condition (13.27) assumes the form

$$\mathbf{a}_r^\dagger \mathbf{T} \mathbf{a}_s = \delta_{rs}$$

where $\mathbf{a}_r$ is the vector corresponding to frequency ω_r.

It is possible to show mathematically that the eigenfrequencies are real although we have given a physical argument to this effect.

As an illustration, let us derive the normalised vectors obtained in the earlier example. On applying condition (13.27), we get equation

$$1 = \tilde{\mathbf{a}}_1 \mathbf{T}_1 \mathbf{a}_1 = ml^2 a_{11}^2 (1 \quad 1) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2ml^2 a_{11}^2$$

or

$$a_{11} = \frac{1}{\sqrt{2ml^2}}$$

Similarly

$$1 = \tilde{\mathbf{a}}_2 \mathbf{T} \mathbf{a}_2 = ml^2 a_{22}^2 (1 \quad -1) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 2ml^2 a_{22}^2$$

or

$$a_{22} = \frac{1}{\sqrt{2ml^2}}$$

Hence, the solutions are from equation (13.15)

$$q_1(t) = \frac{1}{\sqrt{2ml^2}} [\cos(\omega_1 t - \delta_1) + \cos(\omega_2 t - \delta_2)]$$

and

$$q_2(t) = \frac{1}{\sqrt{2ml^2}} [\cos(\omega_1 t - \delta_1) - \cos(\omega_2 t - \delta_2)]$$

13.4 NORMAL COORDINATES

The general solution for q_j has already been shown to be

$$q_j(t) = \sum_r a_{jr} \exp\{i(\omega_r t - \delta_r)\} = \sum_r a_{jr} \cos(\omega_r t - \delta_r)$$

Since the a_{jr} are normalised, there remains no ambiguity in the solution for the q_j. This would mean that it is no longer possible to specify an arbitrary displacement for a particle. But such a restriction is not physically meaningful. To do away with the loss of generality that is introduced on account of the arbitrary normalisation, we introduce a scale factor α

and write

$$q_j(t) = \sum_r \alpha a_{jr} \exp\{i(\omega_r t - \delta_r)\} \tag{13.28}$$

The notation may be simplified by writing $\beta_r = \alpha \exp(-i\delta_r)$. Then, we get

$$q_j(t) = \sum_r \beta_r a_{jr} \exp(i\omega_r t) \tag{13.29}$$

i.e.

$$q_j(t) = \sum_r a_{jr}\eta_r(t) \tag{13.30}$$

where

$$\eta_r(t) = \beta_r \exp(i\omega_r t) \tag{13.31}$$

Equation (13.30) can be written as

$$\mathbf{q} = \mathbf{A}\boldsymbol{\eta} \tag{13.32}$$

where matrix $\mathbf{A} = (a_{jr})$.

From equation (13.31), it is observed that η_r are the quantities that would undergo oscillations with only one frequency ω_r. Hence, these quantities must be considered as new coordinates called the normal coordinates of the system.

In terms of new coordinates, the potential and kinetic energy expressions take a simple form.

$$\begin{aligned} V &= \tfrac{1}{2}\sum_{jk} V_{jk} q_j q_k = \tfrac{1}{2}\sum_{jk} V_{jk}\Big(\sum_r a_{jr}\eta_r\Big)\Big(\sum_s a_{ks}\eta_s\Big) \\ &= \tfrac{1}{2}\sum_{rs}\Big(\sum_{jk} V_{jk} a_{jr} a_{ks}\Big)\eta_r\eta_s = \tfrac{1}{2}\sum_{rs}\omega_r^2\Big(\sum_{jk} T_{jk} a_{jr} a_{ks}\Big)\eta_r\eta_s \\ &= \tfrac{1}{2}\sum_r \omega_r^2\eta_r^2 \end{aligned}$$

or

$$V = \tfrac{1}{2}\tilde{\boldsymbol{\eta}}(\omega_r)^2\boldsymbol{\eta}$$

where we have used equations (13.21) and (13.27). Similarly

$$\begin{aligned} T &= \tfrac{1}{2}\sum_{jk} T_{jk}\dot{q}_j\dot{q}_k = \tfrac{1}{2}\sum_{jk} T_{jk}\Big(\sum_r a_{jr}\dot{\eta}_r\Big)\Big(\sum_s a_{ks}\dot{\eta}_s\Big) \\ &= \tfrac{1}{2}\sum_{rs}\Big(\sum_{jk} T_{jk} a_{jr} a_{ks}\Big)\dot{\eta}_r\dot{\eta}_s = \tfrac{1}{2}\sum_r \dot{\eta}_r^2 \end{aligned}$$

or

$$T = \tfrac{1}{2}\tilde{\dot{\boldsymbol{\eta}}}\dot{\boldsymbol{\eta}}$$

Thus, the Lagrangian becomes

$$L = T - V = \tfrac{1}{2}\sum_r (\dot{\eta}_r^2 - \omega_r^2\eta_r)$$

and we get the equations of motion

$$\frac{\partial L}{\partial \eta_r} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\eta}_r} = 0$$

or

$$\ddot{\eta}_r - \omega_r^2\eta_r = 0 \tag{13.33}$$

Equations (13.33) are a set of n independent equations. Thus equations of motion (13.10) expressed in normal coordinates become completely separable.

In the above derivation, it is revealed that when normal coordinates are used, the potential and the kinetic energies become simultaneously diagonal. It should be remembered that the off-diagonal elements of

tensors **T** and **V** give rise to the coupling between the particles. Thus, when the normal coordinates are used, the tensors become diagonal and hence this will uncouple the coordinates and the problem will become completely separable into the independent motions of the normal coordinates. Each will be characterised by its particular normal frequency.

In order to express the normal coordinates in terms of coordinates q_j we evaluate complex quantities β_r. Let us write

$$\beta_r = \mu_r + i\nu_r \tag{13.34}$$

Then, we have

$$q_j(t) = \sum_r a_{jr}(\mu_r + i\nu_r) \exp(i\omega_r t) \tag{13.35}$$

and

$$\dot{q}_j(t) = \sum_r i\omega_r a_{jr}(\mu_r + i\nu_r) \exp(i\omega_r t) \tag{13.36}$$

From the real part of equation (13.35), we have at $t = 0$

$$q_j(0) = \sum_r \mu_r a_{jr}$$

Multiplying both sides of this equation by $T_{jk}a_{ks}$ and summing over j and k, we obtain

$$\begin{aligned}\sum_{jk} T_{jk}a_{ks}q_j(0) &= \sum_r \mu_r\Big(\sum_{jk} T_{jk}a_{ks}a_{jr}\Big)\\ &= \sum_r \mu_r\delta_{rs} = \mu_r \end{aligned}\tag{13.37}$$

since $\sum_{jk} T_{jk}a_{ks}a_{jr} = \delta_{rs}$.

In a similar manner, considering the real part of equation (13.36) at $t = 0$ and following the same procedure as above, we get

$$\nu_r = -\frac{1}{\omega_s}\sum_{jk} m_{jk}a_{ks}\dot{q}_j(0) \tag{13.38}$$

Using these values of ν_r and μ_r, we can write the normal coordinates as the real part of the expression in $\eta_r = \beta_r \exp(i\omega_r t)$ as

$$\eta_r = \sum_{jk} T_{jk}a_{kr} \exp(i\omega_r t)\left\{q_j(0) - \frac{i}{\omega_r}\dot{q}_j(0)\right\} \tag{13.39}$$

Thus, for arbitrary values of $q_j(0)$ and $\dot{q}_j(0)$, it is possible to find a set of coordinates η_r. Each of these coordinates varies harmonically with single frequency ω_r. The expressions for η_r are, in general, complicated.

Suppose that the coordinates are displaced from their equilibrium positions and are released from there at $t = 0$. Then, at $t = 0$, $q_j(0) \neq 0$ but $\dot{q}_j(0) = 0$.

Under these conditions

$$\eta_r \doteq \exp(i\omega_r t)\sum_{jk} T_{jk}a_{kr}q_j(0),\quad \dot{q}_j(0) = 0 \tag{13.40}$$

Since only the real part of the expression is significant

$$\eta_r = \cos\omega_r t \sum_{jk} T_{jk}a_{kr}q_j(0),\quad \dot{q}_j(0) = 0 \tag{13.41}$$

Using the solutions for the coupled pendulum with the conditions that

positions is imparted to the system as potential energy and is

$$V = \frac{F}{2l}\left[q_1^2 + (q_2 - q_1)^2 + (q_3 - q_2)^2 + \ldots + (q_{n-1} - q_n)^2 + q_n^2\right]$$

$$= \frac{F}{2l}\sum_{j=1}^{n+1}(q_{j-1} - q_j)^2$$

Thus, the Lagrangian of the system is

$$L = \sum_{j=1}^{n+1}\left[\tfrac{1}{2}m\dot{q}_j^2 - \frac{F}{2l}(q_{j-1} - q_j)^2\right]$$

Using Lagrange's equations for q_j, we get

$$m\ddot{q}_j - \frac{F}{l}(q_{j-1} - 2q_j + q_{j+1}) = 0$$

or

$$\ddot{q}_j = \omega_0^2(q_{j-1} - 2q_j + q_{j+1})$$

where we have put $\frac{F}{ml} = \omega_0^2$. The equation of motion for the particle clearly shows that it depends on displacements q_{j-1} and q_{j+1} of its neighbouring particles. This is, therefore, an example of interaction with the 'nearest neighbour' in which coupling is between neighbouring particles only.

Substituting solutions $q_j = a_j \exp(i\omega t)$, we get the set of equations corresponding to equation (13.14)

$$(2\omega_0^2 - \omega^2)a_1 - \omega_0^2 a_2 = 0$$

$$-\omega_0^2 a_1 + (2\omega_0^2 - \omega^2)a_2 - \omega_0^2 a_3 = 0$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$-\omega_0^2 a_{n-1} + (2\omega_0^2 - \omega^2)a_n = 0$$

It is comparatively easy to solve the secular equation for small values of n (Fig. 13.3). Thus, for $n = 1$, we have only one normal mode with frequency $\omega^2 = 2\omega_0^2$. For $n = 2$, the characteristic equation is

$$\begin{vmatrix} 2\omega_0^2 - \omega^2 & -\omega_0^2 \\ -\omega_0^2 & 2\omega_0^2 - \omega^2 \end{vmatrix} = (2\omega_0^2 - \omega^2)^2 - \omega_0^4 = 0$$

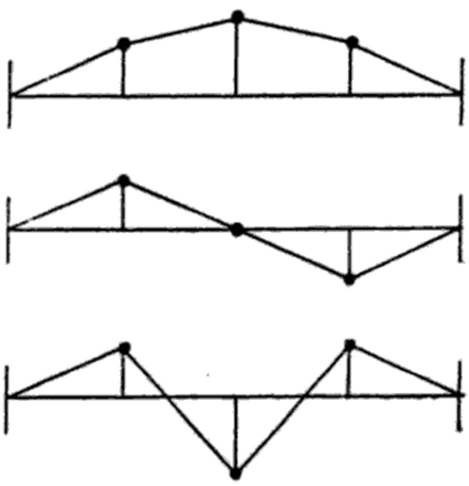


Fig. 13.3 Different modes of vibrations

and we have two normal modes

$$\omega_1^2 = \omega_0^2 \text{ with } a_1 = a_2$$

and
$$\omega_2^2 = 3\omega_0^2 \text{ with } a_1 = -a_2$$

For $n = 3$, the secular equation is

$$\begin{vmatrix} 2\omega_0^2 - \omega^2 & -\omega_0^2 & 0 \\ -\omega_0^2 & 2\omega_0^2 - \omega^2 & -\omega_0^2 \\ 0 & -\omega_0^2 & 2\omega_0^2 - \omega^2 \end{vmatrix} = (2\omega_0^2 - \omega^2)^3 - 2\omega_0^4(2\omega_0^2 - \omega^2) = 0$$

The roots of the equations are $2\omega_0^2$ and $(2 \pm \sqrt{2})\omega_0^2$. Hence, the corresponding normal modes are

$$\omega_1^2 = (2 - \sqrt{2})\omega_0^2, \quad \frac{a_1}{a_2} = \frac{1}{\sqrt{2}} = \frac{a_3}{a_2}$$

$$\omega_2^2 = 2\omega_0^2, \quad \frac{a_1}{a_3} = -1,\ a_2 = 0$$

$$\omega_3^2 = (2 + \sqrt{2})\omega_0^2, \quad \frac{a_1}{a_2} = -\frac{1}{\sqrt{2}} = \frac{a_3}{a_2}$$

In the same way modes for $n = 4, 5, \ldots$ can be found out. It should be noted that for any value of n, the slowest mode is the one in which all the particles are oscillating in the same direction and the fastest mode has the adjacent particles oscillating in opposite directions. For higher value of n, the normal modes approach those of a continuous stretched string.

QUESTIONS

1. For small oscillations we write potential energy V as a quadratic function of displacements and neglect the linear and higher order terms. Explain why.
2. What is meant by stable, unstable and neutral equilibrium? Give examples of each.
3. Show that the eigenvectors corresponding to the two distinct eigenfrequencies are orthogonal. Explain the meaning of orthogonality.
4. The condition of orthonormality of the two vectors expressed by equation (13.27) contains elements T_{jk}. It is different from the usual definition in which T_{jk}'s are absent. How will you interpret the difference?
5. Show that the eigenvalues of symmetric matrix are real.
6. Explain the geometrical meaning of the process of diagonalization.
7. What is degeneracy? Take a few examples of simple systems like double pendulum, linear triatomic molecule, etc. and, without any derivations, discuss their normal modes of vibration and degeneracy. Is degeneracy related with the symmetry of the system? Explain.

14
Special Theory of Relativity

The theory of relativity put forward by Einstein in the year 1905 revolutionised physical concepts during the twentieth century. We shall present a brief discussion of the theory of relativity here. It is well known that physics is a 'science of measurements' of various physical quantities. The theory of relativity reveals that the measurements depend upon the state of motion of the observer as well as upon the quantities that are being measured. The idea of relativity incorporated into mechanics gives rise to relativistic mechanics—in which we come across some peculiar phenomena, particularly when the particles forming a system are moving with high velocity comparable to that of light. The theory of relativity enables us to understand the high energy phenomena in the microscopic as well as macroscopic world.

14.1 NEWTONIAN RELATIVITY

The phenomenon of 'motion' has been of great interest since ages. Although it was recognised that motion of a body involves its displacement relative to something or the other, Newton argued that 'absolute motion is the translation of a body from one absolute place to another absolute place'. But what is meant by 'absolute place' ? The question was never answered. This difficulty was anticipated and Newton stated explicitly that 'a translatory motion can be detected only in the form of a motion relative to other material bodies'.

Motion involves the passage of time. According to Newton, 'absolute, true and mathematical time, of itself and by its own nature, flows uniformly on, without regard to anything external'. Thus, a single time scale would be valid everywhere.

The kind of relativity mentioned in the above statements is called the Newtonian Relativity. We shall now state these views in a mathematical language. Let us represent the position of a point in terms of position

vector $\mathbf{r} = \mathbf{r}(x, y, z)$ and let t denote the time. Then, $\mathbf{r}$ and t specify the position and the instant at which some kind of event occurs. To specify position vector $\mathbf{r}$, however, we require some material reference frame. Similarly to specify instant t we must have some reference process such as the motion of the earth. The instants can be expressed by stating the stages to which the reference process is advanced. The material means used to specify the positions and the instants together with the methods adopted for using them are said to constitute a space-time frame of reference.

We have remarked earlier that an inertial frame of reference is one in which Newton's law of motion, viz., $\mathbf{F} = m\mathbf{a}$ is invariant. Thus, any frame fixed in space or any other frame moving with uniform velocity with respect to the former is an inertial frame of reference.

In Chapter 3, we have seen that if S and S' are two frames of reference in uniform translatory motion with respect to each other, we get transformation equations

$$\mathbf{r}' = \mathbf{r} - \mathbf{v}t \tag{14.1}$$

and

$$t = t'$$

The velocities and the accelerations are given by

$$\dot{\mathbf{r}}' = \dot{\mathbf{r}} - \mathbf{v} \tag{14.2}$$

and

$$\ddot{\mathbf{r}}' = \ddot{\mathbf{r}} \tag{14.3}$$

respectively.

However, from equation (14.2), it is clear that the velocity of a particle is different in the two different systems. The transformations expressed in equations (14.1) are called the Galilean transformations.

Suppose that a source of light is placed at the origin of system S. Let $\mathbf{r}$ represent the position vector of a point on the wave surface which is spherical. The velocity of light is then expressed as

$$\dot{\mathbf{r}} = c\hat{\mathbf{e}}_r \tag{14.4}$$

where $\hat{\mathbf{e}}_r$ is a unit vector in the direction of $\mathbf{r}$. The velocity of light with respect to system S' would be

$$\dot{\mathbf{r}}' = c\hat{\mathbf{e}}_r - \mathbf{v} \tag{14.5}$$

Equation (14.5) shows that the magnitude of the velocity of light is not equal to c. Moreover, it depends upon the direction of propagation of light. Hence, the wave surfaces will not be spherical. A series of experiments were carried out to verify this conclusion. The most crucial of these is the experiment performed by Michelson and Morley in the year 1887.

14.2 MICHELSON-MORLEY EXPERIMENT

The purpose of the Michelson-Morley experiment can be very briefly stated as follows: The light waves propagate through ether with a speed of 3×10^8 m/sec. The ether is a hypothetical medium supposed to be pre-

sent everywhere in order to explain the propagation of light from one point to other. The earth moves through it with a linear speed of about 3×10^4 m/s, and along with the earth we move through ether as a result of the earth's orbital motion around the sun. If the sun is also in motion, the speed with which we are moving through ether may be still greater. If, therefore, two beams of light are sent out such that one travels in the direction of motion of the earth and the other at right angles to this direction, these would require different times for round-trip journeys through the same distance. If this time-difference could be measured, we can determine the velocity of earth with respect to ether, i.e., we can detect the motion of the earth through ether by performing an optical experiment on the earth itself.

In this experiment, a Michelson interferometer (Fig. 14.1) was used.

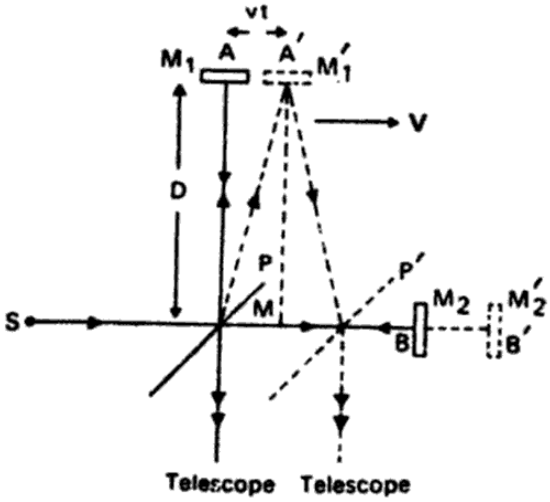


Fig. 14.1 The Michelson-Morley experiment

A beam of light from source S is incident on a semi-silvered glass plate P placed inclined at angle 45° with the incident beam. Each wave of the incident beam is split up into two waves having the same amplitude. One of these two waves is reflected by glass plate P and the other is transmitted through glass plate P. The reflected wave travels towards mirror M_1 and is reflected back by it towards glass plate P. A part of this wave is then transmitted through glass plate P. This part then enters telescope T. The other part of the incident beam, viz., the transmitted wave travels towards mirror M_2, which reflects it back towards glass plate P. A part of this wave is then partially reflected into telescope T. In the actual arrangement of the interferometer, another glass plate (not shown in the figure) is introduced in the path of the transmitted beam so as to make the paths of the two beams equal in glass medium. The two beams enter the telescope along coincident paths and interfere with each other. An interference pattern consisting of bright and dark interference fringes is obtained.

Let both mirrors M_1 and M_2 be at the same distance D from glass plate P. Then, if the apparatus is stationary in ether, the two waves take the same time to return to glass plate P and hence meet in the same phase at the glass plate and also in the telescope.

But, the earth and, hence, the whole apparatus is moving with the same velocity $\mathbf{v}$. Suppose that the direction of motion coincides with the direction of the incident beam (Fig. 14.1). Then, the incident beam strikes glass plate when it is in the position P shown in Fig. 14.1. Due to the motion of the apparatus, the paths of the reflected and the transmitted waves and their reflections at mirrors M_1 and M_2 and at glass plate P will be as shown by the dotted lines. It is obvious that the time required by the two waves for their 'round-trip' journeys through the same distance will not be equal.

The transmitted wave travels towards M_2 approaching it with relative velocity $c - v$. After reflection at M_2, it travels towards glass plate P with velocity $c + v$. Hence, the time required by this wave for its round-trip is given by

$$t_2 = \frac{D}{c - v} + \frac{D}{c + v} = \frac{2Dc}{c^2 - v^2}$$

or

$$t_2 \simeq \frac{2D}{c}\left(1 + \frac{v^2}{c^2}\right) \tag{14.6}$$

The time required by the reflected wave for its round-trip is given by

$$t_1 = 2t' \tag{14.7}$$

where t' is the time required by the wave to travel from P to A'. From triangle $PA'M$, we get

$$c^2t'^2 = v^2t'^2 + D^2$$

Hence

$$t'^2 = \frac{D^2}{c^2 - v^2}$$

or

$$t' = \frac{D}{(c^2 - v^2)^{1/2}} \simeq \frac{D}{c}\left(1 + \frac{v^2}{2c^2}\right) \tag{14.8}$$

Hence

$$t_1 = 2t' = \frac{2D}{c}\left(1 + \frac{v^2}{2c^2}\right) \tag{14.9}$$

From equations (14.6) and (14.9), we find that the difference in the time required by the two waves for their round trips is given by

$$t_2 - t_1 = \frac{Dv^2}{c^3} \tag{14.10}$$

Now, a time difference of one period (T) would amount to a path difference of one wavelength (λ), which will, in turn, amount to a displacement of one fringe across a particular point (crosswire) in the field of view of the telescope. Hence, a time difference $(t_2 - t_1) = \dfrac{Dv^2}{c^3}$ causes $\dfrac{Dv^2}{c^3T} = \dfrac{Dv^2}{c^2\lambda}$ fringes to be displaced across the crosswire. Here, we have used the relation $cT = \lambda$ for light. The displacement of the fringes, however, cannot be noticed since the apparatus is at rest with respect to the obser-

ver all the time. Hence, with an apparatus stationary on earth, no information regarding the time difference between the two paths can be obtained. The apparatus is then slowly turned through 90° so that the two beams interchange their paths. This is equivalent to introducing an additional path difference slowly which would make the displacement of the fringes noticeable. It is obvious that now $\frac{2Dv^2}{c^2\lambda}$ fringes should be displaced.

In the Michelson-Morley experiment effective distance D was made equal to 11 metres by using multiple reflections in mirrors. The wavelength of light used can be taken as 5.9×10^{-7} m, the orbital velocity of the earth as 3×10^4 m/s and the velocity of light as 3×10^8 m/s.

Then, $\frac{2Dv^2}{c^2\lambda} = 0.37$. Thus, 0.37 fringe will be displaced across the crosswire. Michelson and Morley were confident, however, that displacement equal to even a hundredth part of a fringe could be detected with the help of the apparatus. The observed displacements were extremely small as compared to the theoretical value (0.37 of a fringe) and moreover these were not consistent. The experiment was performed in different seasons of the year and at different places. Some more experiments of other kinds also were tried for the same purpose. But, the conclusions were always identical. Thus, the motion of the earth with respect to ether cannot be detected.

The negative result of the Michelson-Morley experiment led to the following conclusions: (i) the hypothesis of existence of ether was rendered untenable by demonstrating that ether has no measurable properties, (ii) the speed of light in free space is constant irrespective of the motion of the source or the observer.

14.3 SPECIAL THEORY OF RELATIVITY

We saw in the previous article that the motion of the earth through ether cannot be detected by performing an optical experiment on the earth itself. It was also recognised that no experiment in mechanics, electricity or magnetism would help us in detecting the motion of the earth through ether. Ether played the part of a universal frame of reference with respect to which waves of light were supposed to propagate. But, when the very existence of ether was thought to be doubtful as suggested by the negative result of the Michelson-Morley experiment, the possibility of having a universal frame of reference was ruled out. The theory of relativity is the outcome of the analysis of the physical consequences that would result if no universal frame of reference is in existence. The Special Theory of Relativity, developed by Einstein in the year 1905, deals with the problems involving the inertial frames of reference. By an inertial frame of reference, we mean the frame of reference in which the law of inertia holds. They move with uniform velocity with respect to one another. The General Theory of Relativity,

proposed by Einstein in the year 1915, deals with the problems involving frames of reference accelerated with respect to an inertial frame.

The special theory of relativity is based on the following postulates:

(i) *The laws of physics should be expressed in equations having the same form in all frames of reference moving with a uniform velocity with respect to one another.*

(ii) *The speed of light in free space has the same value for all the observers irrespective of their state of motion.*

The first postulate brings out the fact that a universal frame of reference is not in existence. If the laws of physics were different for observers in a relative uniform motion with respect to one another, these differences would enable us to determine which objects are 'stationary' in space and which objects are 'moving'. But such a universal frame of reference is not available and we are unable to notice such differences in physical laws. Hence, the postulate.

The second postulate follows directly from the negative results of Michelson-Morley experiment and many other experiments performed for the same purpose.

14.4 LORENTZ TRANSFORMATIONS

In article 14.1, we have stated an equation which represents the relation between the position vectors of a point in the systems S and S' respectively, the systems being in a uniform relative motion with respect to each other (Fig. 14.2). If the motion of system S' with respect to S is along x-axis only, the equation

$$\mathbf{r}' = \mathbf{r} - \mathbf{v}t$$

can be written in its component form, viz.

$$x' = x - vt \tag{14.11}$$

$$y' = y \tag{14.12}$$

and

$$z' = z \tag{14.13}$$

Further, since we do not experience anything contrary, we assume

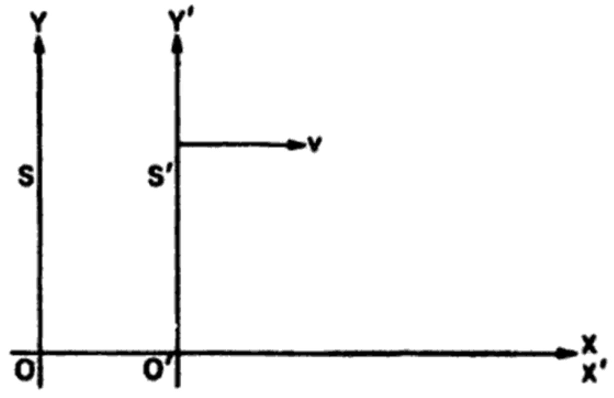


Fig. 14.2 Frame of reference S' moving with velocity v relative to S

that

$$t' = t \tag{14.14}$$

Equations (14.11) to (14.14) are called the Galilean transformations.

If a velocity is measured in system S and we wish to get the components of velocity in system S' in terms of those in system S, we can write

$$V'_x = \frac{dx'}{dt} = V_x - v \tag{14.15}$$

$$V'_y = \frac{dy'}{dt} = V_y \tag{14.16}$$

and

$$V'_z = \frac{dz'}{dt} = V_z \tag{14.17}$$

It is clear that the Galilean transformations and the velocity transformations obtained in equations (14.15) to (14.17), violate the postulates of the special theory of relativity. The equations of physics obtained by these transformations assume totally different forms in frames S and S'. Moreover, if the velocity of light in frame S is c, that in system S' would be $c' = c - v$. It is, therefore, necessary that different transformation equations must be obtained if the postulates of the special theory of relativity are to be satisfied. These transformation equations—called the Lorentz transformations—are developed as follows:

Let us assume that a possible relation between x and x' is of the type

$$x' = k(x - vt) \tag{14.18}$$

where k is a constant of proportionality. It does not depend upon x and t but it may be a function of v. In choosing the relation between x and x' in the form expressed in equation (14.18) the following points are considered: (i) The relation is linear in x and x'. Hence, a single event in frame S corresponds to a single event in S'. (ii) It can be easily reduced to the form $x' = x - vt$ which is known to be correct in Newtonian mechanics if $k = 1$ when $v \ll c$.

The inverse relation will be written as

$$x = k'(x' + vt') \tag{14.19}$$

in which primed quantities are replaced by unprimed quantities and *vice versa* and v by $-v$. The other relations will be as before

$$y' = y \tag{14.20}$$

and

$$z' = z \tag{14.21}$$

Times t and t' are not equal. To verify this, let us substitute the value of x' from equation (14.18) in equation (14.19). This gives

$$x = kk'(x - vt) + k'vt'$$

or

$$t' = kt + \left(\frac{1 - kk'}{k'v}\right)x \tag{14.22}$$

To determine the value of k and k' we use the second postulate of the special theory of relativity.

Let at $t = 0$, the origins of the two systems S and S' coincide with each other. Let this instant correspond to $t' = 0$ as well. Suppose that a signal of light is given out from the common origin of S and S' at $t = t' = 0$. The signal propagates in the two systems satisfying the equations

$$x = ct \tag{14.23}$$

and

$$x' = ct' \tag{14.24}$$

in systems S and S' respectively.

Substituting the values of x' and t' from equations (14.18) and (14.22) into equation (14.24), we obtain

$$k(x - vt) = ckt + \left(\frac{1 - kk'}{k'v}\right)cx$$

Hence

$$x = ct\left[\frac{1 + v/c}{1 - \left\{\frac{1}{kk'} - 1\right\}\frac{c}{v}}\right] \tag{14.25}$$

Comparing equation (14.25) with equation (14.23), we get

$$\frac{1 + v/c}{1 - \left\{\frac{1}{kk'} - 1\right\}\frac{c}{v}} = 1$$

or

$$\sqrt{kk'} = \frac{1}{\sqrt{1 - v^2/c^2}}$$

Now, kk' depends on v^2 and not on v, the relative velocity of S' with respect to S. In fact, we cannot choose between the two frames S and S' except for the sign of v which does not affect the dependence of kk'. Therefore, we choose

$$k = k' = \frac{1}{\sqrt{1 - v^2/c^2}} \tag{12.26}$$

Substituting these values in equations (14.18) and (14.22), we obtain

$$\left.\begin{aligned} x' &= \frac{x - vt}{\sqrt{1 - v^2/c^2}} \\ y' &= y \\ z' &= z \\ \text{and} \qquad t' &= \frac{t - \dfrac{vx}{c^2}}{\sqrt{1 - v^2/c^2}} \end{aligned}\right\} \tag{14.27}$$

Equations (14.27) are called the Lorentz transformations.

The inverse Lorentz transformations can now be written as

$$\left.\begin{aligned} x &= \frac{x' + vt'}{\sqrt{1 - v^2/c^2}} \\ y &= y' \end{aligned}\right\} \tag{14.28}$$

observer in frame S' is at rest with respect to S' and hence with respect to the rod or the rod is at rest with respect to this observer. Length L_0 of the rod measured by him is expressed as

$$L_0 = x_2' - x_1' \tag{14.32}$$

where x_2' and x_1' are the coordinates of the extremities of the rod. Thus, L_0 is the length of the rod in the frame of reference in which the rod is at rest.

Now, let us find length L of the rod in system S relative to which the rod is in motion with velocity v. The Lorentz transformation equations give

$$x_1' = \frac{x_1 - vt_1}{\sqrt{1 - v^2/c^2}} \tag{14.33}$$

and

$$x_2' = \frac{x_2 - vt_2}{\sqrt{1 - v^2/c^2}} \tag{14.34}$$

Hence, we get

$$L_0 = x_2' - x_1' = \frac{x_2 - x_1 - v(t_2 - t_1)}{\sqrt{1 - v^2/c^2}} \tag{14.35}$$

But, $x_2 - x_1 = L$, the length of the rod measured in frame S. Here, t_2 and t_1 are the times at which the end coordinates x_2 and x_1 of the rod are measured. Since the measurement should be simultaneous in frame S for determining the length of the rod, we have $t_2 = t_1$. Therefore

$$L_0 = \frac{L}{\sqrt{1 - v^2/c^2}}$$

or

$$L = L_0\sqrt{1 - v^2/c^2} \tag{14.36}$$

i.e., the length of the rod in motion with respect to an observer appears to the observer to be shorter than when it is at rest with respect to him. The phenomenon is known as the Lorentz-FitzGerald contraction.

Since the relative velocity of the two frames S and S' appears as v^2 in equation (14.36), the opinion of observers is reciprocal. Thus, in general, it can be stated that the length of an object in motion relative to an observer appears to him to be shorter than the length when measured at rest. It is, therefore, clear that the length of an object is a maximum in a frame of reference in which it is stationary.

It should be noted that the effects of length contraction become significant only when the velocity of the objects approaches the velocity of light. Thus, if $v = 0.9c$, ratio $\frac{L}{L_0}$ is about 0.44. Further, the effect occurs only in the direction of relative motion. If the relative velocity is parallel to the x-axis, the y and z dimensions of the object are unaffected.

An interesting consequence of the length contraction was furnished by Terrel in 1959. Consider a cube of each side L_0 which moves with uniform velocity **v** with respect to an observer situated at some distance. The direction of motion of the cube is perpendicular to the line of sight

Let L be the distance travelled by the meson in its own frame of reference moving at $0.998c$ before decaying. Thus, $L = 600$ m. Distance L_0 corresponds to our frame of reference and is given by

$$L_0 = \frac{L}{\sqrt{1 - v^2/c^2}}$$

$$= \frac{600}{\sqrt{1 - (0.998)^2}} \simeq 9500 \text{ m}$$

Hence, in spite of the short lifetime, the μ-mesons are able to reach the ground from great altitudes at which they are created.

(c) Time Dilation

Consider a clock situated at position x' in moving frame S'. Suppose that t_1' and t_2' are two instants recorded by the observer in frame S'. Then, the time interval measured by him is given by

$$t_0 = t_2' - t_1' \tag{14.37}$$

The observer in frame S, however, measures these instants as

$$t_1 = \frac{t_1' + \dfrac{vx'}{c^2}}{\sqrt{1 - v^2/c^2}} \tag{14.38}$$

and

$$t_2 = \frac{t_2' + \dfrac{vx'}{c^2}}{\sqrt{1 - v^2/c^2}} \tag{14.39}$$

respectively.

Thus, the time interval according to the observer in frame S is

$$t = t_2 - t_1$$

$$= \frac{t_2' - t_1'}{\sqrt{1 - v^2/c^2}}$$

or

$$t = \frac{t_0}{\sqrt{1 - v^2/c^2}} \tag{14.40}$$

Since, $\sqrt{1 - \dfrac{v^2}{c^2}}$ is a fraction, we conclude that a clock measures a longer time interval between events occurring in its own frame than the time interval measured by a clock in a frame moving relative to it. In other words, to an observer in motion relative to the clock, the time intervals appear to be lengthened. This phenomenon is called *time dilation*. Since, equation (14.40) involves factor v^2, the opinion of the observers will be reciprocal.

Although time is a relative quantity, we are able to observe the following phenomena: (i) Time does not run backward for any observer. The sequence of events in a series of events is never altered for any observer. At the most, the intervals of time between any two events may be different.

This is because the velocity available for communication is always less than or equal to the velocity of light c. (ii) No observer can see an event before it takes place. In other words, there is no way of probing into the future.

14.6 ADDITION OF VELOCITIES

Consider a body that moves with respect to both frames of reference S and S'. An observer in frame S measures the components of velocity $\mathbf{V}$ of the body as

$$V_x = \frac{dx}{dt},\ V_y = \frac{dy}{dt} \text{ and } V_z = \frac{dz}{dt} \tag{14.41}$$

An observer belonging to the frame of reference S' makes the following measurements of velocity $\mathbf{V}'$ of the body measured by him.

$$V'_x = \frac{dx'}{dt'},\ V'_y = \frac{dy'}{dt'} \text{ and } V'_z = \frac{dz'}{dt'} \tag{14.42}$$

From the Lorentz-transformations [equation (14.27)], we can write

$$dx' = \frac{dx - v\,dt}{\sqrt{1 - v^2/c^2}}$$

$$dy' = dy$$

$$dz' = dz$$

and

$$dt' = \frac{dt - \dfrac{v\,dx}{c^2}}{\sqrt{1 - v^2/c^2}} \tag{14.43}$$

Thus

$$V'_x = \frac{dx'}{dt'} = \frac{dx - v\,dt}{dt - v\dfrac{dx}{c^2}}$$

$$= \frac{\dfrac{dx}{dt} - v}{1 - \dfrac{v}{c^2}\dfrac{dx}{dt}}$$

or

$$V'_x = \frac{V_x - v}{1 - \dfrac{v}{c^2}V_x} \tag{14.44}$$

The y-components of the velocities will be related by the equation

$$V'_y = \frac{dy'}{dt'}$$

$$= \frac{dy}{\left(dt - v\dfrac{dx}{c^2}\right)\Big/\sqrt{1 - v^2/c^2}}$$

$$= \frac{dy\sqrt{1 - v^2/c^2}}{dt - v\dfrac{dx}{c^2}}$$

$$= \frac{V_y\sqrt{1-v^2/c^2}}{1-\dfrac{vV_x}{c^2}} \tag{14.45}$$

Similarly, we can write the relation between z-components of the velocities as

$$V_z' = \frac{V_z\sqrt{1-v^2/c^2}}{1-\dfrac{vV_x}{c^2}} \tag{14.46}$$

Equations (14.44) to (14.46) constitute the relativistic transformation equations for velocity. If the relative velocity of S' with respect to S, i.e., v is negligible in comparison with the velocity of light, these equations reduce to classical expressions stated in equations (14.15) to (14.17).

The inverse transformation relations of equations (14.44), (14.45) and (14.46) are

$$V_x = \frac{V_x' + v}{1+\dfrac{vV_x}{c^2}} \tag{14.47}$$

$$V_y = \frac{V_y'\sqrt{1-v^2/c^2}}{1+\dfrac{vV_x'}{c^2}} \tag{14.48}$$

and

$$V_z = \frac{V_z'\sqrt{1-v^2/c^2}}{1+\dfrac{vV_x'}{c^2}} \tag{14.49}$$

Suppose that $V_x' = c$. It means that a ray of light is emitted in the moving frame of reference S' in the same direction as that of its motion with respect to system S. Then, an observer in system S would measure this velocity as

$$V_x = \frac{V_x' + c}{1+\dfrac{vV_x'}{c^2}}$$

$$= \frac{c+v}{1+\dfrac{vc}{c^2}} = c \tag{14.50}$$

As could be expected this is consistent with the second postulate of the theory of relativity. We would arrive at the same conclusion if we start with $V_x = c$ and use equation (14.44) to calculate V_x'. It, therefore, appears that c is the highest limit to the velocity that can be acquired by material bodies.

Suppose that a rocket is moving with a speed of $0.9c$ with respect to the earth in a certain direction. We wish to overtake it with a speed which is more than $0.9c$ by $0.4c$. Then, according to classical physics, we must move with velocity $1.3c$ with respect to the earth. Thus, our speed should be greater than that of light. This contradicts the statement

made earlier. The theory of relativity provides the clue to the problem. Let $v = 0.9c$ be the velocity of the rocket (S') with respect to the earth (S). Then, $V'_x = 0.4c$, the velocity of our rocket with respect to S'. Naturally, our velocity with respect to earth is given by

$$V_x = \frac{V'_x + v}{1 + \frac{vV'_x}{c^2}}$$

$$= \frac{0.4c + 0.9c}{1 + \frac{0.4c \times 0.9c}{c^2}}$$

$$= \frac{1.3}{1.36}c = 0.9779c \tag{14.51}$$

which is certainly less than the velocity of light.

14.7 VARIATION OF MASS WITH VELOCITY

In non-relativistic mechanics, the inertial mass m given in, say, $\mathbf{p} = m\mathbf{v}$ is constant. However, the same might not be true in relativistic mechanics. We should, however, expect that the laws of relativistic mechanics should reduce to those of classical mechanics when the velocity of the particles is very small as compared to the velocity of light ($v \ll c$). We shall now assume that the mass depends on its velocity, i.e., $m \equiv m(u)$ and when $u = 0$, $m =$ the rest-mass of the particle. To determine the dependence of mass on velocity, we shall consider collision of two bodies and further assume that (i) the law of conservation of momentum, and (ii) the conservation of the relativistic masses of the particles are valid.

Consider two bodies each of mass m' moving in opposite directions along the x'-axis with velocities u' and $-u'$ as observed from frame of references S' (Fig. 14.4). Let these bodies collide and coalesce into one body. The body thus formed will be at rest according to the law of conservation of momentum with respect to system S'.

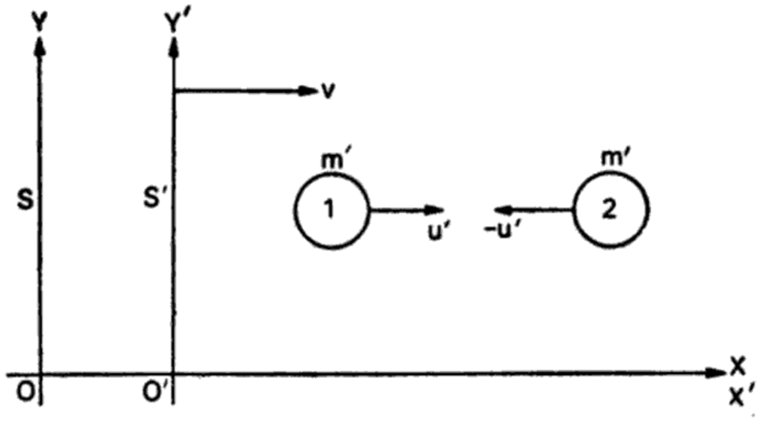


Fig. 14.4 Two bodies 1 and 2, each of mass m', moving with equal and opposite velocities along the x'-axis in frame S'

If the collision of the two bodies is observed from frame of reference

S, the velocities of the two bodies as observed from S will be given by

$$u_1 = \frac{u' + v}{1 + \frac{u'v}{c^2}} \quad \text{and} \quad u_2 = \frac{-u' + v}{1 - \frac{u'v}{c^2}} \tag{14.52}$$

where u_1 and u_2 are the velocities along the x-axis. Here, we have used the relativistic law for addition of velocities. Let m_1 and m_2 be the masses of the two bodies with respect to frame S. Then, the body formed when the two bodies coalesce into each other has a mass $(m_1 + m_2)$ by the law of conservation of mass and it moves with a velocity v along the x-axis with respect to S. Note that this body is at rest with respect to S'. Then, by the law of conservation of momentum, we can write

$$m_1u_1 + m_2u_2 = (m_1 + m_2)v \tag{14.53}$$

i.e.
$$m_1\left[\frac{u' + v}{1 + \frac{u'v}{c^2}}\right] + m_2\left[\frac{-u' + v}{1 - \frac{u'v}{c^2}}\right] = (m_1 + m_2)v \tag{14.54}$$

Dividing equation (14.54) throughout by m_2 and simplifying, we get

$$\frac{m_1}{m_2} = \frac{1 + \frac{u'v}{c^2}}{1 - \frac{u'v}{c^2}} \tag{14.55}$$

Now consider the factor $\left(1 - \frac{u_1^2}{c^2}\right)$. Substituting the value of u_1 as given in equation (14.52), we get

$$1 - \frac{u_1^2}{c^2} = 1 - \frac{1}{c^2}\left(\frac{u' + v}{1 + \frac{u'v}{c^2}}\right)^2$$

$$= \frac{c^2\left(1 + \frac{u'v}{c^2}\right)^2 - (u' + v)^2}{c^2\left(1 + \frac{u'v}{c^2}\right)^2}$$

$$= \frac{\left(1 - \frac{v^2}{c^2}\right)(c^2 - u'^2)}{c^2\left(1 + \frac{u'v}{c^2}\right)^2}$$

$$= \frac{\left(1 - \frac{v^2}{c^2}\right)\left(1 - \frac{u'^2}{c^2}\right)}{\left(1 + \frac{u'v}{c^2}\right)^2} \tag{14.56}$$

Hence
$$\sqrt{1 - \frac{u_1^2}{c^2}} = \frac{\sqrt{\left(1 - \frac{v^2}{c^2}\right)\left(1 - \frac{u'^2}{c^2}\right)}}{\left(1 + \frac{u'v}{c^2}\right)}$$

Integrating by parts, we get

$$\begin{aligned}
T &= \frac{m_0 v^2}{\sqrt{1 - v^2/c^2}} - m_0 \int_0^v \frac{v\,dv}{\sqrt{1 - v^2/c^2}} \\
&= \frac{m_0 v^2}{\sqrt{1 - v^2/c^2}} + m_0 c^2 \left[\sqrt{1 - \frac{v^2}{c^2}}\right]_0^v \\
&= \frac{m_0 c^2}{\sqrt{1 - v^2/c^2}} - m_0 c^2 \\
&= mc^2 - m_0 c^2 \\
&= (\Delta m)c^2 \qquad (14.65)
\end{aligned}$$

where $\Delta m = m - m_0$.

Thus, the kinetic energy of a body is equal to the product of the increase in the mass and the square of the speed of light.

Equation (14.65) can be written as

$$mc^2 = T + m_0 c^2 \qquad (14.66)$$

If we call $mc^2 = E$, the total energy of the body, the energy of the body at rest is equal to $E_0 = m_0 c^2$. Quantity E_0 is called the rest (mass) energy of the body. Hence, equation (14.66) can be written as

$$E = E_0 + T \qquad (14.67)$$

The expression for rest energy viz. $E_0 = mc^2$ shows that mass is yet another form of energy. In fact, conversion of matter into energy is a source of energy liberated in all exothermic reactions.

Since mass and energy are related to each other, we have to consider a principle of conservation of mass and energy. Mass can be created or destroyed, provided that an equivalent amount of energy vanishes or is being created and *vice versa*.

The relativistic expression for the kinetic energy

$$T = mc^2 - m_0 c^2$$

can be easily reduced to the classical expression, viz.

$$T = \tfrac{1}{2} m_0 v^2$$

if we make v very small as compared to c.

This simplification is carried out as follows:

$$\begin{aligned}
T &= mc^2 - m_0 c^2 \\
&= \frac{m_0 c^2}{\sqrt{1 - \frac{v^2}{c^2}}} - m_0 c^2 \\
&= m_0 c^2 \left(1 - \frac{v^2}{c^2}\right)^{-1/2} - m_0 c^2 \\
&= m_0 c^2 \left(1 + \frac{v^2}{2c^2} \cdots\right) - m_0 c^2 = \tfrac{1}{2} m_0 v^2
\end{aligned}$$

It will be observed that all formulae of relativity reduce to the corresponding formulae of classical mechanics at low speeds. The relativistic formulation of mechanics is a more accurate approach while

(x_1, x_2, x_3 and x_4) for (x, y, z and ict). The coordinates of point P are related to each other by the relation

$$\begin{pmatrix} x_1' \\ x_4' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_4 \end{pmatrix}$$

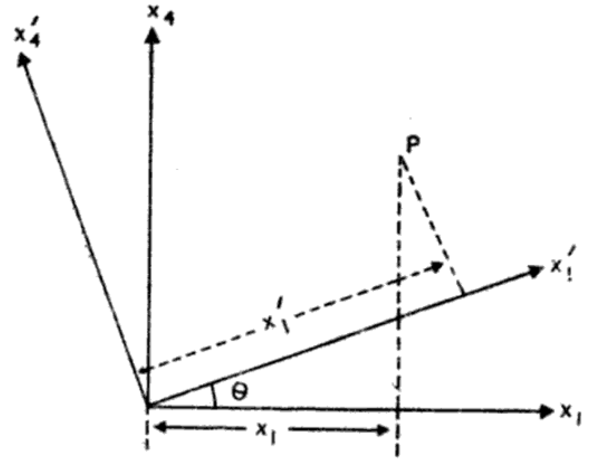


Fig. 14.6 Invariance of s^2

The inverse relation is

$$\begin{pmatrix} x_1 \\ x_4 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1' \\ x_4' \end{pmatrix}$$

If $x_1' = 0$, $\qquad x_1 = -\sin\theta\, x_4'$

and $\qquad x_4 = \cos\theta\, x_4'$

Hence

$$-\tan\theta = \frac{x_1}{x_4} = \frac{x}{ict} = \frac{v}{ic}$$

or

$$\tan\theta = i\frac{v}{c}$$

From this, we get

$$\cos\theta = \frac{1}{\sqrt{1 - v^2/c^2}} \text{ and } \sin\theta = \frac{iv/c}{\sqrt{1 - v^2/c^2}}$$

Thus, the above transformation equations become

$$x_1' = \frac{1}{\sqrt{1 - v^2/c^2}}\left(x_1 + i\frac{v}{c}x_4\right) = \frac{1}{\sqrt{1 - v^2/c^2}}(x_1 - vt)$$

$$x_2' = x_2$$

$$x_3' = x_3$$

$$x_4' = \frac{1}{\sqrt{1 - v^2/c^2}}\left(-i\frac{v}{c}x_1 + x_4\right) = \frac{ic}{\sqrt{1 - v^2/c^2}}\left(-\frac{v}{c^2}x_1 + t\right)$$

These equations are precisely the Lorentz transformations.

The four coordinates x, y, z and ict define a vector in the four-space. Such a vector is often termed a four-vector. We shall give the definitions

of some of the four-vectors in the next article. The four-vector remains fixed in four-space irrespective of any rotation of the coordinate system. Thus, a four-vector is independent of the frame of reference to which it is referred.

Introduction of four-vectors in the four-space has another advantage. When physical laws are expressed in terms of four-vectors then the law takes the same form in any frame moving with a uniform velocity with another frame. Thus, the invariance of the physical law is automatically proved when expressed in four-vectors.

We now use the concept of four-space to understand some implications of the theory of relativity. Consider two events 1 and 2 plotted on rectangular axes x and ct. The straight lines passing through the origin of the axes and satisfying equation $x = \pm ct$ determine cones as shown in Fig. 14.7. These cones are termed as light cones. Let event 1 occur at $x = 0$, $t = 0$ and event 2 at $x = \Delta x$ and $t = \Delta t$.

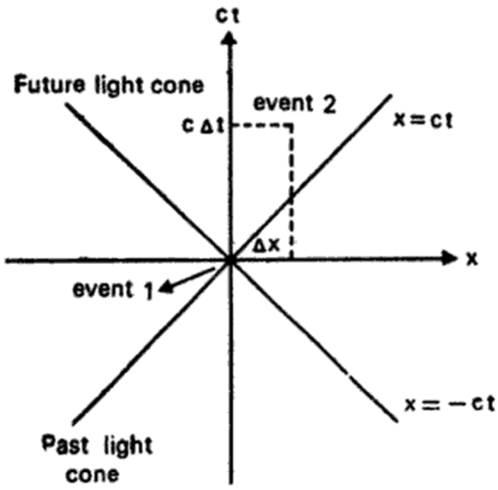


Fig. 14.7 Light cones

The interval Δs between these events is defined by

$$\Delta s^2 = (c\,\Delta t)^2 - (\Delta x)^2 \tag{14.70}$$

It is obvious that $(\Delta s)^2$ is also invariant to the Lorentz transformation. Hence, we can write

$$(\Delta s)^2 = (c\,\Delta t)^2 - (\Delta x)^2 = (c\,\Delta t')^2 - (\Delta x')^2 \tag{14.71}$$

Thus, the conclusions that we draw in frame S at the origin of which is event 1, are also valid in any other frame in relative uniform motion with respect to S.

Event 2 will be related to event 1 in some way or the other if and only if a signal travelling at a speed lower than that of light can connect them. Thus

$$c\,\Delta t > |\Delta x|$$

or

$$(\Delta s)^2 > 0 \tag{14.72}$$

Such an interval is called a time-like interval. All events that could

have affected event 1 lie in the past light cone of event 1. Similarly, all events that are affected by event 1, lie in the future light cone of event 1 (Fig. 14.7).

If in any case, the relation

$$c\,\Delta t < |\Delta x|$$

or

$$(\Delta s)^2 < 0 \tag{14.73}$$

is satisfied, the two points cannot be related by the light signal. In order to relate the events at the two points one would need a signal that would travel faster than light. Since this is unphysical, such events should not be allowed in the physical theory. If a light signal is created at one point it will not be received at the other point considered above. This is expressed by saying that there is no cause and effect relationship, i.e., no causal relationship between the two points. The interval is called the space-like interval. Any event that is related to event 1 by a space-like interval lies outside the light cones of event 1. Such an event has not interacted with event 1 in the past nor would it interact with it in future. The two events are totally independent of each other.

If

$$c\,\Delta t = |\Delta x|$$

or

$$(\Delta s)^2 = 0 \tag{14.74}$$

the two events can be connected by a light signal only and the corresponding interval is called a light-like interval. Events that are connected with event 1 by light-like intervals lie on the boundaries of the light cones of event 1.

Similar conclusions can be drawn by considering event 2 at $x = 0$, $t = 0$ as well. This is because $(\Delta s)^2$ is invariant. For example, if event 2 is inside the future cone of event 1, event 1 must be inside the past cone of event 2 and *vice versa*.

Thus, in general, events that lie in the future of an event as seen from a frame of reference S lie in its future in every other frame S'. A similar statement can be made about the past. Thus, the future and the past have absolute or invariant meanings (Fig. 14.8). But 'simultaneity' is a

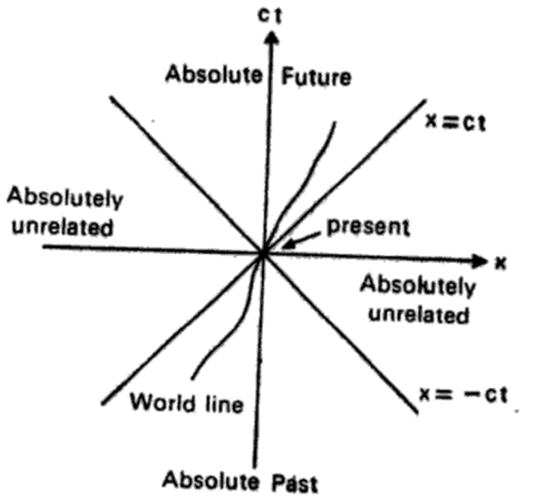


Fig. 14.8 World line of a particle

relative concept and hence another event that lies inside the light cone may appear to occur simultaneously with event 1 in some frame of reference.

The path of a particle in four-space is called a world line. The world line of a particle must lie in its light cones.

14.10 FOUR-VECTORS

In the Minkowski space, a quantity is called a four-vector if it has four components, each of which can be transformed by using the Lorentz transformation equations.

A four-vector **X** which is used to represent the position in a four-space or a four-position vector is written as

$$\mathbf{X}_\mu = (x, y, z, ict) \equiv (x_1, x_2, x_3, x_4)$$

or
$$\mathbf{X}_\mu = (\mathbf{x}, ict) = (\mathbf{x}, x_4) \tag{14.75}$$

wherein the first three, i.e., space components of the four-position vector **X** are the usual three-dimensional position vector **x** and the fourth component is $x_4 = ict$.

The differential of **X** is also a four-vector given by

$$d\mathbf{X}_\mu = (d\mathbf{x}, dx_4 = ic\ dt) \tag{14.76}$$

The length element in the four-space can be given by

$$\begin{aligned} ds &= \sqrt{\sum_{\mu=1}^{4} d\mathbf{X}_\mu^2} \\ &= \sqrt{\sum_{j=1}^{3} dx_j^2 - c^2\, dt^2} \end{aligned} \tag{14.77}$$

We use the usual notation to denote values 1 to 4 by the Greek letter μ and 1 to 3 values in real space by the Latin letter j. Length ds is invariant to the Lorentz transformations.

Now we define a quantity

$$d\tau = \sqrt{dt^2 - \frac{1}{c^2}\sum_{j=1}^{3} dx_j^2} = \frac{i}{c}\sqrt{\sum_{\mu=1}^{4} d\mathbf{X}_\mu^2} \tag{14.78}$$

which is found to be invariant to the Lorentz transformations since it is $\frac{i}{c}\,ds$. Quantity $d\tau$ is called the element of proper time in the Minkowski four-space.

The ratio of four-vector $d\mathbf{X}_\mu$ to the invariant proper time $d\tau$ is also a four-vector called the four-velocity **V**. Thus

$$\mathbf{V}_\mu = \frac{d\mathbf{X}_\mu}{d\tau} = \left(\frac{d\mathbf{x}}{dt}, ic\frac{dt}{d\tau}\right) \tag{14.79}$$

The first three components of the velocity in a three-dimensional space are

$$v_j = \frac{dx_j}{dt}, j = 1, 2, 3 \tag{14.80}$$

Hence, we write

$$\mathbf{F} = \frac{d}{dt}\left[\frac{m_0\mathbf{v}}{\sqrt{1 - v^2/c^2}}\right] \tag{14.87}$$

where $\mathbf{F}$ is a three-dimensional force vector.

We have seen above that

$$p_4 = \frac{im_0c}{\sqrt{1 - v^2/c^2}} = imc = i\frac{E}{c} \tag{14.88}$$

where $E = mc^2$ is the total energy of the particle. We can, therefore, write four momentum vector $\mathbf{P}$ as

$$\mathbf{P} = m_0\mathbf{V} = \left(\mathbf{p},\ i\frac{E}{c}\right) \tag{14.89}$$

where $\mathbf{p}$ stands for the three space components of the momentum.

The square of four-velocity $\mathbf{V}$ is given by

$$V^2 = \sum_{\mu=1}^{4} V_\mu^2 = \frac{v^2 - c^2}{1 - v^2/c^2} = -c^2 \tag{14.90}$$

which is found to be invariant.

Hence, the square of the four-momentum will be

$$P^2 = \sum_{\mu=1}^{4} P_\mu^2 = m_0^2V^2 = -m_0^2c^2 \tag{14.91}$$

and is invariant.

But

$$\mathbf{p}\cdot\mathbf{p} = p^2$$

Hence, from equation (14.89)

$$P^2 = p^2 - \frac{E^2}{c^2} \tag{14.92}$$

From equations (14.91) and (14.92), we get

$$E^2 = p^2c^2 + m_0^2c^4 \tag{14.93}$$

a relation which has already been stated.

The gradient operator in a four-space will have components corresponding to $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$ and $\frac{\partial}{ic\,\partial t}$. The D'Alembertian operator in four-space is defined as

$$\begin{aligned}\Box^2 &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \\ &= \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\end{aligned} \tag{14.94}$$

where ∇^2 is the Laplacian operator.

In the equation of continuity

$$\nabla\cdot\mathbf{j} + \frac{d\rho}{dt} = 0 \tag{14.95}$$

wavelength and is denoted by λ_c. Thus

$$\lambda' - \lambda = 2\lambda_c \sin^2 \frac{\theta}{2} \tag{14.106}$$

Equation (14.106) gives the Compton shift in wavelength.

To calculate the energy of the recoil electron, let us first evaluate ν' in terms of ν. For this, we use equation (14.105), viz.

$$\frac{c}{\nu'} - \frac{c}{\nu} = \frac{h}{m_0 c}(1 - \cos\theta)$$

or

$$\frac{1}{\nu'} = \frac{1}{\nu} + \frac{h}{m_0 c^2}(1 - \cos\theta)$$

Hence

$$\nu' = \frac{m_0 c^2}{h\left[(1 - \cos\theta) + \frac{m_0 c^2}{h\nu}\right]} \tag{14.107}$$

Thus

$$\begin{aligned} h\nu' &= \text{energy of the scattered photon} \\ &= \frac{m_0 c^2}{(1 - \cos\theta) + \frac{m_0 c^2}{h\nu}} \\ &= \frac{m_0 c^2}{(1 - \cos\theta) + \frac{1}{\alpha}} \end{aligned} \tag{14.108}$$

where $\alpha = \frac{h\nu}{m_0 c^2}$. It will be clear that

$$h\nu' = \frac{h\nu}{1 + \alpha(1 - \cos\theta)} \tag{14.109}$$

Moreover

$$\begin{aligned} \frac{\nu'}{\nu} &= \frac{m_0 c^2}{h\nu(1 - \cos\theta) + m_0 c^2} \\ &= \frac{1}{\alpha(1 - \cos\theta) + 1} \end{aligned} \tag{14.110}$$

Now, energy T of the recoil electron is given by

$$\begin{aligned} T &= h\nu - h\nu' \\ &= h\nu - \frac{h\nu}{\alpha(1 - \cos\theta) + 1} \\ &= h\nu\left[1 - \frac{1}{2\alpha \sin^2 \frac{\theta}{2} + 1}\right] \\ &= h\nu \cdot \frac{2\alpha \sin^2 \frac{\theta}{2}}{2\alpha \sin^2 \frac{\theta}{2} + 1} \end{aligned}$$

But

$$\alpha = \frac{h\nu}{m_0 c^2} = \frac{h}{m_0 c} \cdot \frac{\nu}{c} = \frac{\lambda_c}{\lambda}$$

where λ_c is the Compton wavelength.

Hence

$$T = h\nu \frac{2\left(\frac{\lambda_c}{\lambda}\right)\sin^2\frac{\theta}{2}}{2\left(\frac{\lambda_c}{\lambda}\right)\sin^2\frac{\theta}{2}+1} \tag{14.111}$$

Finally, the relation between scattering angle θ and angle of recoil φ of the electron is given by

$$\tan\varphi = \frac{h\nu'\sin\theta}{h\nu - h\nu'\cos\theta}$$

from equations (14.100) and (14.101).

Hence

$$\tan\varphi = \frac{\nu'\sin\theta}{\nu - \nu'\cos\theta}$$

But $\nu' = \dfrac{\nu}{2\alpha\sin^2\theta/2+1}$ by equation (14.109).

Hence

$$\tan\varphi = \frac{\nu\sin\theta}{2\alpha\sin^2\theta/2+1}\,\frac{1}{\left[\nu - \dfrac{\nu\cos\theta}{2\alpha\sin^2\theta/2+1}\right]}$$

or

$$\tan\varphi = \frac{1}{(1+\alpha)\tan\theta/2} \tag{14.112}$$

QUESTIONS

1. What displacement is invariant in going from one inertial frame to another? What interval is invariant? Explain the terms invariants and covariants.
2. What would absolute simultaneity imply?
3. How do parallel velocities combine?
4. How does a Lorentz transformation rotate the axes of a Minkowski coordinate system?
5. A rod of proper length L is at rest in the xy plane making an angle θ with the x-axis. What does an observer moving at speed u along the x-axis find for its length and angle of inclination?
6. Can a particle move through a medium with a speed greater than that of light in that medium? Explain.
7. A person sitting in a moving train cannot determine, by an experiment performed in the train, whether the train is moving or not. Comment.
8. The speed of light is the same for all observers regardless of the state of their motion. Explain.

9. Determine the values of the time dilation factor if
$$\frac{v}{c}=\frac{3}{5}, \frac{4}{5} \text{ and } \frac{12}{13}$$
10. Two particles approach each other with a speed of $0.9c$ each with respect to the laboratory frame. Find their relative speed.
11. If two photons approach each other, what is their relative speed according to the Newtonian relativity and according to the special theory of relativity? Which one is correct?
12. Four dimensional volume element is invariant under the Lorentz transformation. Explain.
13. Show that the relativistic expression for kinetic energy tends to the classical expression if $v \ll c$.
14. What is meant by the space-time interval? When is this interval said to be space-like, time-like and light-like?
15. Is the relativistic contraction isotropic? Explain.
16. Are the relativistic and inertial masses equal? Explain.
17. Manufacturers of fast oscilloscopes claim that "the writing speed of electron beams in their oscilloscopes is greater than that of light." Comment.
18. Is the concept of a perfectly rigid body acceptable in relativity? Explain.
19. Draw a neat diagram of light cones indicating past and future. Also show the world line in it.
20. Does **F** equal $m\mathbf{a}$ in relativity? Does $m\mathbf{a}$ equal $\frac{d}{dt}(mu)$ in relativity? Explain.
21. In what frame can all kinetic energy be converted to rest energy?
22. Explain clearly the distinction between the origin of variation of mass of a classical system and that of a relativistic system.

PROBLEMS

1. An observer on the moon observes two space-ships coming towards him from opposite directions at speeds $0.8c$ and $0.85c$ respectively. Find the relative speed of the two spaceships as observed by an observer on any one of them.
2. How fast must an electron move in order that its mass is equal to (a) twice its rest mass, (b) equal to the rest mass of a proton?
3. Find the speed of a 1-MeV electron, both classically and relativistically.
4. Find the amount of mass gained by (a) an electron, and (b) a proton when each is accelerated so that the kinetic energy of each becomes 100 MeV.

of the rod cross certain reference index marks. Show that this definition also leads to the 'length contraction'.

17. An event occurs in frame S at $x = 50$ km, $y = 10$ km and $z = 2$ km at $t = 1 \times 10^{-6}$ s. If frame S' moves with velocity $0.94c$ along the common x-x' axis, the origins coinciding at $t = t' = 0$, find co-ordinates x', y', z' and t' of this event in S'.
18. At what speed v will the Galilean and Lorentz expressions for x differ by 1 per cent? By 15 per cent?
19. What is the proper time interval between the occurrence of two events if in some inertial frame the two events are separated by 10^5 m and occur 4 s apart?
20. An airplane 30 m in length in its rest system is moving with a uniform speed of 600 m/s with respect to the earth. By what fraction of its rest length will it appear to be shortened to an observer on the earth? How long will it take by earth's clocks for the airplane's clock to fall behind by 10^{-6} s?
21. The rest radius of the earth is 6.4×10^6 m and its orbital speed about the sun is about 3×10^4 m/s. By how much would the earth's diameter appear to be shortened to an observer on the sun, due to earth's orbital motion?
22. Derive the relativistic acceleration transformation

$$a'_x = \frac{a_x\left(1 - \dfrac{v^2}{c^2}\right)^{1/2}}{\left(1 - \dfrac{V_x v}{c^2}\right)^3}$$

where $a_x = \dfrac{dV_x}{dt}$ and $a'_x = \dfrac{dV_x}{dt'}$

23. The earth receives radiant energy from the sun at the rate of 1.34×10^3 watts/m^3. At what rate is the sun losing mass due to its radiation? Given: the rest mass of sun $= 2 \times 10^{30}$ kg and mean separation between the earth and the sun $= 1.49 \times 10^{11}$ m.
24. Prove that if $\frac{v}{c} \ll 1$, the kinetic energy of a moving particle will always be much less than its rest energy m_0c^2.
25. Show that a particle which travels at the speed of light must have a zero rest-mass. Also show that for a particle of zero rest-mass, $v = c$, $T = E$ and $p = \dfrac{E}{c}$.
26. Find the effective mass of a photon of a wavelength (a) 5890 Å, and (b) 1.0 Å.
27. The nucleus of C^{12} consists of six protons and six neutrons held together by strong nuclear forces. Given: mass of C^{12} to be 12.000000 a.m.u.; mass of proton to be 1.007825 a.m.u., and mass of neutron to be 1.008665 a.m.u. Find the binding energy of a C^{12} nucleus.

28. A body of mass m at rest breaks up spontaneously into two parts having masses m_1 and m_2 and speeds v_1 and v_2 respectively. Show that $m > m_1 + m_2$ using conservation of mass-energy.

29. A body of rest-mass m_0 and travelling with speed $0.7c$ makes a completely inelastic collision (i.e. coalesce) with an identical body initially at rest. Find the speed of the resulting single body and its rest mass.

30. Show that when $\frac{v}{c} < \frac{1}{10}$, the classical expressions for kinetic energy, viz. $T = \frac{1}{2}m_0v^2$ and linear momentum, viz. $p = m_0v$, may be used with an error of less than one per cent.

APPENDIX A
Coordinate Systems

So far, we have mainly used the cartesian coordinate system to represent motion of a particle. This system, however, is not always convenient to solve all sorts of problems. Many a time, we come across a problem having certain symmetries which decide the choice of a coordinate system. Thus, the motion of a particle in a central force field, i.e., when $\mathbf{F} = \hat{\mathbf{e}}_r F(\mathbf{r})$ has a spherical symmetry and can be conveniently studied if plane polar coordinates are used. The motion of a charged particle moving along a spiral path in a magnetic field has an axis of symmetry and can be conveniently treated in a cyclindrical polar coordinate system. Quite a large number of coordinate systems are developed and the reader is referred to the books on mathematical physics given in the list of references. We shall first consider the curvilinear coordinates and restrict ourselves only to the spherical polar and cylindrical coordinate systems.

A.1 CURVILINEAR COORDINATES

In the cartesian coordinate system, the intersections of two of the three planes $x = 0$, $y = 0$ and $z = 0$ taken in order are used to define z, x and y axes respectively (Fig. A.1).

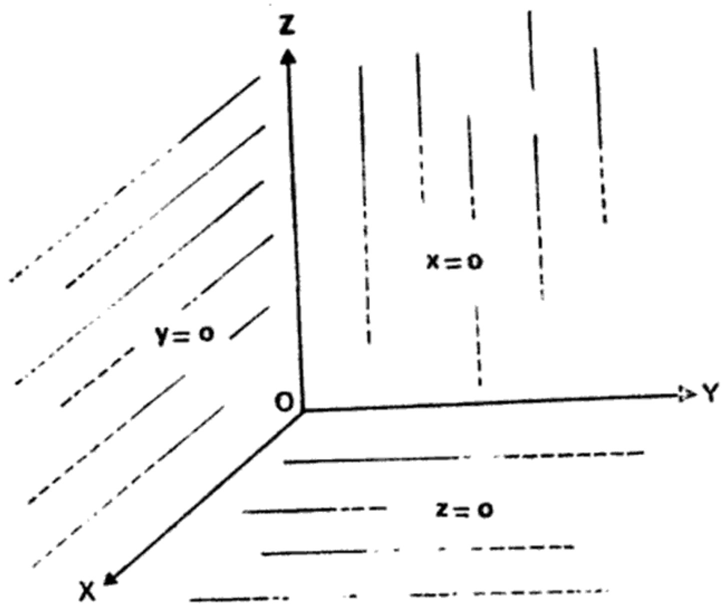


Fig. A.1 The cartesian coordinate system

In general, we need not restrict ourselves to the plane surfaces to define the coordinate axes. Any suitable set of three curved surfaces can be used as reference surfaces and their intersections as the reference axes. Such a system is called a curvilinear system.

The axes will, in general, be curved (Fig. A.2). Let us denote the

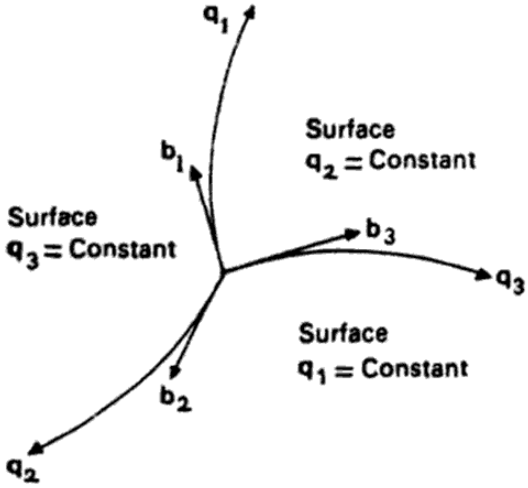


Fig. A.2 The curvilinear coordinates and generalised base vectors

curved coordinate axes by q_1, q_2 and q_3 respectively. It should be noted that the q_1-axis is an intersection of two surfaces $q_2 =$ constant and $q_3 =$ constant and so on. Cartesian coordinates x, y and z are related to q_1, q_2 and q_3 by the relations which can be expressed as

$$\begin{aligned} x &= x(q_1, q_2, q_3) \\ y &= y(q_1, q_2, q_3) \qquad \text{(A.1)} \\ \text{and} \quad z &= z(q_1, q_2, q_3) \end{aligned}$$

Exact relationship can be written only when we know expressions for surfaces.

Equation (A.1) gives the transformation equations from one coordinate system to another. The inverse transformation equations can be written as

$$\left.\begin{aligned} q_1 &= q_1(x, y, z) \\ q_2 &= q_2(x, y, z) \\ \text{and} \quad q_3 &= q_3(x, y, z) \end{aligned}\right\} \qquad \text{(A.2)}$$

In order to understand the transformations, consider a simple example of plane polar coordinates (Fig. A.3). From this it is clear that

$$x = r\cos\theta$$

and

$$y = r\sin\theta$$

The inverse transformation equations are

$$r = \sqrt{x^2 + y^2}$$

and

$$\theta = \tan^{-1}\frac{y}{x}$$

We now define the base vectors. Base vector $\mathbf{b}_i$ is drawn in the direction of change of position vector $\mathbf{r}$ produced by an infinitesimal change in coordinate q_i when the other coordinates are fixed.

Thus
$$\mathbf{b}_i = \frac{\partial \mathbf{r}}{\partial q_i}, \quad i = 1, 2, 3 \tag{A.3a}$$

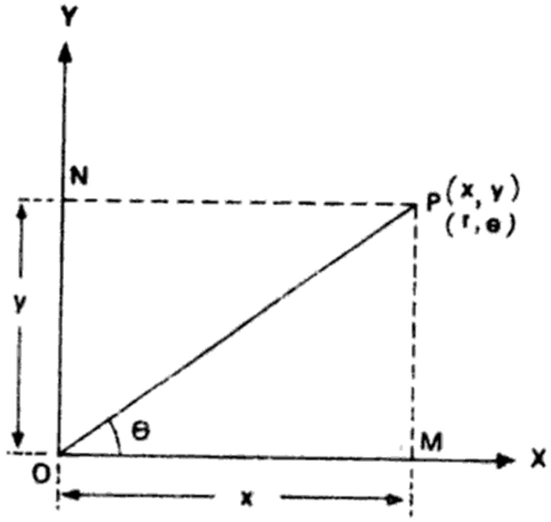


Fig. A.3 The plane polar coordinates and the cartesian coordinates

In general, a base vector need not be a unit vector. In fact it can be related to unit vector $\hat{\mathbf{e}}_i$ by the formula

$$\mathbf{b}_i = h_i \hat{\mathbf{e}}_i \tag{A.3b}$$

where
$$h_i = \left|\frac{\partial \mathbf{r}}{\partial q_i}\right|$$

From equation (A.3b), we can write

$$\hat{\mathbf{e}}_i = \frac{1}{h_i}\mathbf{b}_i$$

or
$$\hat{\mathbf{e}}_i = \frac{\partial \mathbf{r}/\partial q_i}{|\partial \mathbf{r}/\partial q_i|}, \quad i = 1, 2, 3 \tag{A.4}$$

In terms of unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ of the cartesian coordinate system, from equations (A.2) and (A.3a) we can write

$$\mathbf{b}_i = \mathbf{i}\frac{\partial x}{\partial q_i} + \mathbf{j}\frac{\partial y}{\partial q_i} + \mathbf{k}\frac{\partial z}{\partial q_i} = h_i \hat{\mathbf{e}}_i \tag{A.5}$$

Since $\hat{\mathbf{e}}_i$ is a unit vector, we have from equation (A.5)

$$h_i = \sqrt{\left(\frac{\partial x}{\partial q_i}\right)^2 + \left(\frac{\partial y}{\partial q_i}\right)^2 + \left(\frac{\partial z}{\partial q_i}\right)^2} \tag{A.6}$$

In order to define the coordinate system, the three surfaces $q_1 =$ constant, $q_2 =$ constant and $q_3 =$ constant must obviously be non-coplanar; this

means that the three base vectors must be non-coplanar, condition for which is

$$(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = \begin{vmatrix} \dfrac{\partial x}{\partial q_1} & \dfrac{\partial y}{\partial q_1} & \dfrac{\partial z}{\partial q_1} \\ \dfrac{\partial x}{\partial q_2} & \dfrac{\partial y}{\partial q_2} & \dfrac{\partial z}{\partial q_2} \\ \dfrac{\partial x}{\partial q_3} & \dfrac{\partial y}{\partial q_3} & \dfrac{\partial z}{\partial q_3} \end{vmatrix} \neq 0 \tag{A.7}$$

This determinant is known as the Jacobian and the condition of equation (A.7) is put in a compact notation as

$$\frac{\partial(x, y, z)}{\partial(q_1, q_2, q_3)} \neq 0 \tag{A.8}$$

The distance between two neighbouring points in the cartesian coordinate system is given by

$$dr^2 = dx^2 + dy^2 + dz^2 \tag{A.9}$$

But, differential increments dx, dy and dz in x, y and z, respectively, are related to the increments in q_1, q_2 and q_3 through the usual rules of partial differentiation as

$$\left.\begin{aligned} dx &= \frac{\partial x}{\partial q_1} dq_1 + \frac{\partial x}{\partial q_2} dq_2 + \frac{\partial x}{\partial q_3} dq_3 \\ dy &= \frac{\partial y}{\partial q_1} dq_1 + \frac{\partial y}{\partial q_2} dq_2 + \frac{\partial y}{\partial q_3} dq_3 \\ \text{and} \qquad dz &= \frac{\partial z}{\partial q_1} dq_1 + \frac{\partial z}{\partial q_2} dq_2 + \frac{\partial z}{\partial q_3} dq_3 \end{aligned}\right\} \tag{A.10}$$

Substituting the values from equation (A.10) into equation (A.9), we get

$$dr^2 = \sum_{ij} h_{ij}\, dq_i\, dq_j \tag{A.11}$$

where we have introduced a symbol

$$h_{ij} = \frac{\partial x}{\partial q_i}\frac{\partial x}{\partial q_j} + \frac{\partial y}{\partial q_i}\frac{\partial y}{\partial q_j} + \frac{\partial z}{\partial q_i}\frac{\partial z}{\partial q_j} \tag{A.12}$$

The aggregate of coefficients h_{ij} is called a metric and it specifies the nature of the coordinate system. From equations (A.11) and (A.12), it is clear that, for cartesian coordinates

$$h_{ij} = 0 \quad \text{for } i \neq j$$

and

$$h_{ii} = 1 \quad \text{for } i = 1, 2, 3$$

The curvilinear coordinate system considered so far is non-orthogonal, i.e., base vectors $\mathbf{b}_i$ are not mutually perpendicular to one another. We normally use orthogonal coordinate systems and hence shall restrict our attention to such systems alone.

A.2 ORTHOGONAL CURVILINEAR COORDINATES

For the axes of a coordinate system to be orthogonal, the unit vectors

along these must satisfy the conditions

$$\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij}$$

and

$$h_{ij} = 0 \quad \text{for} \quad i \neq j \tag{A.13}$$

Then, the distance between two neighbouring points is given by

$$dr^2 = (h_1\, dq_1)^2 + (h_2\, dq_2)^2 + (h_3\, dq_3)^2 \tag{A.14}$$

where $h_1 = \sqrt{h_{11}}$, $h_2 = \sqrt{h_{22}}$ and $h_3 = \sqrt{h_{33}}$ from equations (A.12) and (A.6). Thus, dr^2 is the sum of squares of the three components along the three axes. The cross-product terms of the type $h_{ij}q_iq_j$ are absent. In fact, this is the advantage of choosing an orthogonal coordinate system. Equation (A.14) suggests that displacement vector $d\mathbf{r}$ can be written as

$$d\mathbf{r} = \mathbf{b}_1\, dq_1 + \mathbf{b}_2\, dq_2 + \mathbf{b}_3\, dq_3$$

or

$$d\mathbf{r} = \hat{\mathbf{e}}_1 h_1\, dq_1 + \hat{\mathbf{e}}_2 h_2\, dq_2 + \hat{\mathbf{e}}_3 h_3\, dq_3 \tag{A.15}$$

The three components of $d\mathbf{r}$ are written as

$$dr_i = h_i\, dq_i, \qquad i = 1, 2, 3 \tag{A.16}$$

or

$$d\mathbf{r}_i = \hat{\mathbf{e}}_i h_i\, dq_i, \qquad i = 1, 2, 3 \tag{A.17}$$

It should be noted that q_i may not necessarily have the dimensions of length. Quantities h_1, h_2 and h_3 appearing in the above equations as the multipliers of increments in coordinates are called the scale factors. The scale factors depend, in general, on q_i and may have dimensions.

A.3 ELEMENT OF SURFACE AREA

We have already considered that the plane area is a vector quantity, and is expressed as the cross product of two vectors $d\mathbf{r}_1$ and $d\mathbf{r}_2$. Hence, the element of area is given by

$$\begin{aligned} |d\mathbf{r}_1 \times d\mathbf{r}_2| &= |\hat{\mathbf{e}}_1 \times \hat{\mathbf{e}}_2|\, h_1 h_2\, dq_1\, dq_2 \\ &= h_1 h_2\, dq_1\, dq_2 \\ &= d\sigma_{12} \end{aligned}$$

or

$$d\mathbf{r}_1 \times d\mathbf{r}_2 = d\boldsymbol{\sigma}_{12} \tag{A.18}$$

Thus, three components of the differential vector-area element can, in general, be written as

$$|d\boldsymbol{\sigma}_k| = d\sigma_{ij} = h_i h_j\, dq_i\, dq_j \tag{A.19}$$

where i, j and k must be taken in cyclic order.

A.4 VOLUME ELEMENT

The scalar triple product of three vectors represents the volume of the parallelopiped formed by the three vectors as its sides. Let $d\mathbf{r}_1$, $d\mathbf{r}_2$ and $d\mathbf{r}_3$ be three differential length elements; then the volume element is given by

$$\begin{aligned} d\tau &= d\mathbf{r}_1 \cdot d\mathbf{r}_2 \times d\mathbf{r}_3 \\ &= (\hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_2 \times \hat{\mathbf{e}}_3)\, h_1 h_2 h_3\, dq_1\, dq_2\, dq_3 \\ &= h_1 h_2 h_3\, dq_1\, dq_2\, dq_3 \end{aligned} \tag{A.20}$$

since the coordinate system is orthogonal.

In articles 2.5, 2.6 and 2.8, we have obtained expressions for the gradient, the divergence and the curl in differential form in the cartesian coordinate system. The cartesian coordinates and the other orthogonal coordinate systems can be considered to be special cases of orthogonal curvilinear coordinates. Hence, it is instructive and useful to obtain general expressions for the gradient, the divergence and the curl in such a system.

A.5 GRADIENT IN ORTHOGONAL CURVILINEAR COORDINATES

We have defined the gradient of a scalar point function as a vector having a magnitude equal to the maximum space rate of change of the scalar function and drawn in a direction in which this maximum space rate of change of scalar function is obtained. We extend the same idea to define the gradient in terms of the orthogonal curvilinear coordinates.

Let $\Phi = \Phi(q_1, q_2, q_3)$ be a scalar function.

Then, the component of grad Φ in the direction of normal to surface q_1 = a constant is given by

$$\nabla\Phi\,|_1 = \frac{\partial\Phi}{\partial r_1} = \frac{\partial\Phi}{h_1\,\partial q_1} \tag{A.21}$$

This component is directed along vector $\hat{\mathbf{e}}_1$ or $\mathbf{b}_1$. Similar expression for $\nabla\Phi|_2$ and $\nabla\Phi|_3$ can be written. Combining all these expressions vectorially, we get

$$\nabla\Phi = \hat{\mathbf{e}}_1\frac{\partial\Phi}{h_1\,\partial q_1} + \hat{\mathbf{e}}_2\frac{\partial\Phi}{h_2\,\partial q_2} + \hat{\mathbf{e}}_3\frac{\partial\Phi}{h_3\,\partial q_3} \tag{A.22}$$

From equation (A.22), we obtain the vector differential operator as

$$\nabla = \hat{\mathbf{e}}_1\frac{\partial}{h_1\,\partial q_1} + \hat{\mathbf{e}}_2\frac{\partial}{h_2\,\partial q_2} + \hat{\mathbf{e}}_3\frac{\partial}{h_3\,\partial q_3} \tag{A.23}$$

This operator assumes the familiar form in the cartesian coordinates, if we substitute

$$\hat{\mathbf{e}}_1 = \mathbf{i},\ \hat{\mathbf{e}}_2 = \mathbf{j} \text{ and } \hat{\mathbf{e}}_3 = \mathbf{k}$$

$$h_1 = h_2 = h_3 = 1$$

and $$q_1 = x,\ q_2 = y \text{ and } q_3 = z$$

A.6 DIVERGENCE IN ORTHOGONAL CURVILINEAR COORDINATES

The divergence of a vector function $\mathbf{V}$ has been defined as the limiting value of the surface integral of vector $\mathbf{V}$ over the surface bounding the volume element $d\tau$ per unit volume as the volume of the element tends to zero. Thus

$$\nabla\cdot\mathbf{V} = \lim_{d\tau\to 0}\frac{\int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma}}{d\tau} \tag{A.24}$$

Consider a volume $d\tau = h_1h_2h_3\,dq_1\,dq_2\,dq_3$ as shown in Fig. A.4. In order to find $\nabla\cdot\mathbf{V}$, we have to compute the surface integral of vector $\mathbf{V}$

over the surfaces of elementary volume $d\tau$, viz. $\int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma}$.

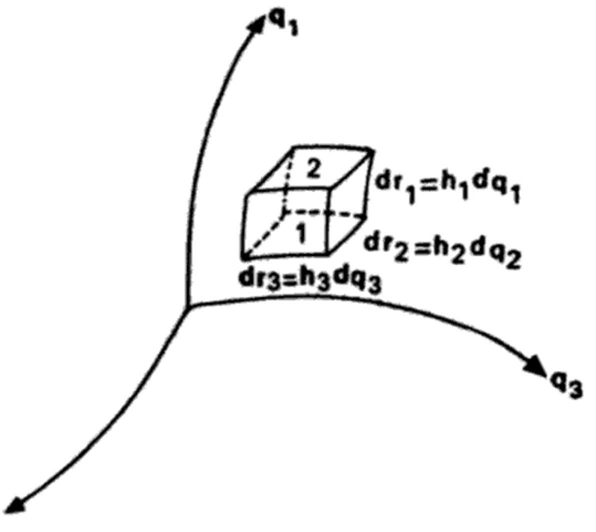


Fig. A.4 Divergence of a vector in curvilinear coordinates

First, consider surfaces marked 1 and 2. Over these surfaces, q_1 will remain constant at certain values. Consider vector **V** which has average values V_1 and $V_1 + \delta V_1$ over the surfaces 1 and 2 respectively. Hence, the contribution of these surfaces towards the surface integral is

$$\begin{aligned}(V_1 + \delta V_1)h_2h_3\, dq_2\, dq_3 &- V_1h_2h_3\, dq_2\, dq_3 \\ &= \delta V_1 h_2h_3\, dq_2\, dq_3 \\ &= \frac{\partial V_1}{\partial q_1} h_2h_3\, dq_1\, dq_2\, dq_3 \\ &= \frac{1}{h_1h_2h_3}\frac{\partial(V_1h_2h_3)}{\partial q_1}\, d\tau \end{aligned} \tag{A.25}$$

Similarly, we can find the contributions of other pairs of opposite surfaces towards the surface integral. Adding all these contributions, we get

$$\int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma} = \frac{1}{h_1h_2h_3}\left[\frac{\partial(V_1h_2h_3)}{\partial q_1} + \frac{\partial(V_2h_3h_1)}{\partial q_2} + \frac{\partial(V_3h_1h_2)}{\partial q_3}\right] d\tau \tag{A.26}$$

Hence, the divergence of vector **V** is given by

$$\begin{aligned}\nabla\cdot\mathbf{V} &= \lim_{d\tau\to 0}\frac{\int_\sigma \mathbf{V}\cdot d\boldsymbol{\sigma}}{d\tau} \\ &= \frac{1}{h_1h_2h_3}\left[\frac{\partial(V_1h_2h_3)}{\partial q_1} + \frac{\partial(V_2h_3h_1)}{\partial q_2} + \frac{\partial(V_3h_1h_2)}{\partial q_3}\right]\end{aligned} \tag{A.27}$$

The expression for the Laplacian operator can be at once obtained by substituting $\mathbf{V} = \nabla\Phi$, where $\Phi = \Phi(q_1, q_2, q_3)$ is a scalar function. Thus

$$\nabla\cdot\mathbf{V} = \nabla\cdot\nabla\Phi = \nabla^2\Phi = \frac{1}{h_1h_2h_3}\left[\frac{\partial\left(\frac{h_2h_3}{h_1}\frac{\partial\Phi}{\partial q_1}\right)}{\partial q_1} + \frac{\partial\left(\frac{h_3h_1}{h_2}\frac{\partial\Phi}{\partial q_2}\right)}{\partial q_2} + \frac{\partial\left(\frac{h_1h_2}{h_3}\frac{\partial\Phi}{\partial q_3}\right)}{\partial q_3}\right] \tag{A.28}$$

A.7 CURL OF A VECTOR IN ORTHOGONAL CURVILINEAR COORDINATES

The curl of a vector function **V** has already been defined as the limiting maximum value of the line integral of the vector over the periphery of an element of surface area per unit area as it tends to zero. It is a vector directed along the positive normal to the plane of the area when it is in the position of giving maximum value.

Thus
$$\nabla \times \mathbf{V} \cdot \hat{\mathbf{n}} = \lim_{d\sigma \to 0} \frac{\oint \mathbf{V} \cdot d\mathbf{l}}{d\sigma} \tag{A.29}$$

We now apply this definition to find out the component of curl **V** along the q_3-direction (Fig. A.5). Consider an element of area $d\sigma = h_1 h_2\, dq_1\, dq_2$ described on a surface $q_3 =$ constant. Its direction will be along the q_3-axis if it is described in the sense shown in Fig. A.5.

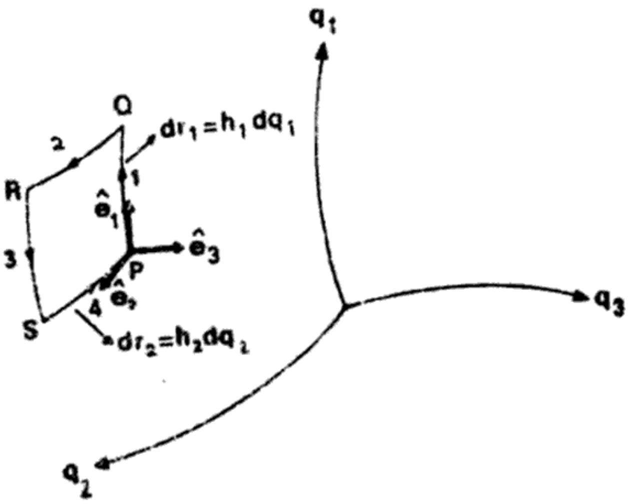


Fig. A.5 Curl of a vector in curvilinear coordinates

We want to evaluate the line integral along the path denoted by *PQRSP*. Let V_1 and V_2 be the components of **V** along the q_1 and q_2 axes at point P. Let $\frac{\partial V_1}{\partial q_2}$ and $\frac{\partial V_2}{\partial q_1}$ be the rates of changes of V_1 and V_2 with respect to q_2 and q_1 respectively.

Then, the values of the components along the paths *PQ*, *QR*, *SR* and *PS* are V_1, $V_2 + \frac{\partial V_2}{\partial q_1} dq_1$, $V_1 + \frac{\partial V_1}{\partial q_2} dq_2$ and V_2 respectively.

Hence, the line integral along the path *PQRSP* is given by

$$\oint \mathbf{V} \cdot d\mathbf{l} = V_1 h_1\, dq_1 + \left(V_2 + \frac{\partial V_2}{\partial q_1} dq_1\right) h_2\, dq_2 - \left(V_1 + \frac{\partial V_1}{\partial q_2} dq_2\right) h_1\, dq_1 - V_2 h_2\, dq_2$$

$$= \left[\frac{\partial}{\partial q_1}(V_2 h_2) - \frac{\partial}{\partial q_2}(V_1 h_1)\right] dq_1\, dq_2 \tag{A.30}$$

It should be noted that the positive sign is used for those paths which are parallel to the component of $\mathbf{V}$ and the negative sign when they are antiparallel.

Equation (A.30) when divided by the area of $PQRS$, viz. $d\sigma_3 = h_1h_2\, dq_1\, dq_2$ represents, in the limit as $d\sigma_3 \to 0$, the component of $\nabla \times \mathbf{V}$ in the q_3-direction.

Thus
$$\nabla \times \mathbf{V}\,|_3 = \frac{1}{h_1h_2}\left[\frac{\partial}{\partial q_1}(V_2h_2) - \frac{\partial}{\partial q_2}(V_1h_1)\right] \tag{A.31}$$

The other two components can be obtained in a similar manner. Adding all the components vectorially, we can write

$$\nabla \times \mathbf{V} = \frac{1}{h_1h_2h_3}\begin{vmatrix} h_1\hat{\mathbf{e}}_1 & h_2\hat{\mathbf{e}}_2 & h_3\hat{\mathbf{e}}_3 \\ \dfrac{\partial}{\partial q_1} & \dfrac{\partial}{\partial q_2} & \dfrac{\partial}{\partial q_3} \\ h_1V_1 & h_2V_2 & h_3V_3 \end{vmatrix} \tag{A.32}$$

A.8 RECTANGULAR CARTESIAN COORDINATES

This is the simplest and a familiar coordinate system in which the coordinates are $q_1 = x$, $q_2 = y$ and $q_3 = z$.

The unit vectors are $\hat{\mathbf{e}}_1 = \mathbf{i}$, $\hat{\mathbf{e}}_2 = \mathbf{j}$ and $\hat{\mathbf{e}}_3 = \mathbf{k}$. These are constant vectors. The scale factors are given by

$$h_1 = h_x = 1,\ h_2 = h_y = 1 \text{ and } h_3 = h_z = 1.$$

A.9 SPHERICAL POLAR COORDINATES

In this coordinate system, the position of a point is fixed by coordinates

$$q_1 = r,\ q_2 = \theta \text{ and } q_3 = \varphi$$

From Fig. A.6, it will be clearly seen that spherical polar coordinates r, θ and φ are related to cartesian coordinates x, y and z by the relations

$$\begin{aligned} x &= r \sin\theta \cos\varphi \\ y &= r \sin\theta \sin\varphi \\ \text{and} \quad z &= r\cos\theta \end{aligned} \tag{A.33}$$

From equations (A.33) it can be shown that

(i) $r = \sqrt{x^2 + y^2 + z^2}$ = constant describes a family of concentric spheres with the centre at the origin,

(ii) $\theta = \cos^{-1}\dfrac{z}{\sqrt{x^2+y^2+z^2}}$ = constant describes right-circular cones with vertex at the origin, and

(iii) $\varphi = \tan^{-1}\dfrac{y}{x}$ = constant describes the half planes through the polar axis, viz. the z-axis.

Position vector $\mathbf{r}$ is given by

$$\begin{aligned} \mathbf{r} &= x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \\ &= r\sin\theta\cos\varphi\ \mathbf{i} + r\sin\theta\sin\varphi\ \mathbf{j} + r\cos\theta\ \mathbf{k} \end{aligned} \tag{A.34}$$

The limits for these coordinates are

$$0 \leqslant r \leqslant \infty, \quad 0 \leqslant \theta \leqslant \pi \text{ and } 0 \leqslant \varphi \leqslant 2\pi$$

From equation (A.12), the scale factors are given by

$$h_1 = h_r = 1,\ h_2 = h_\theta = r \text{ and } h_3 = h_\varphi = r \sin\theta \tag{A.35}$$

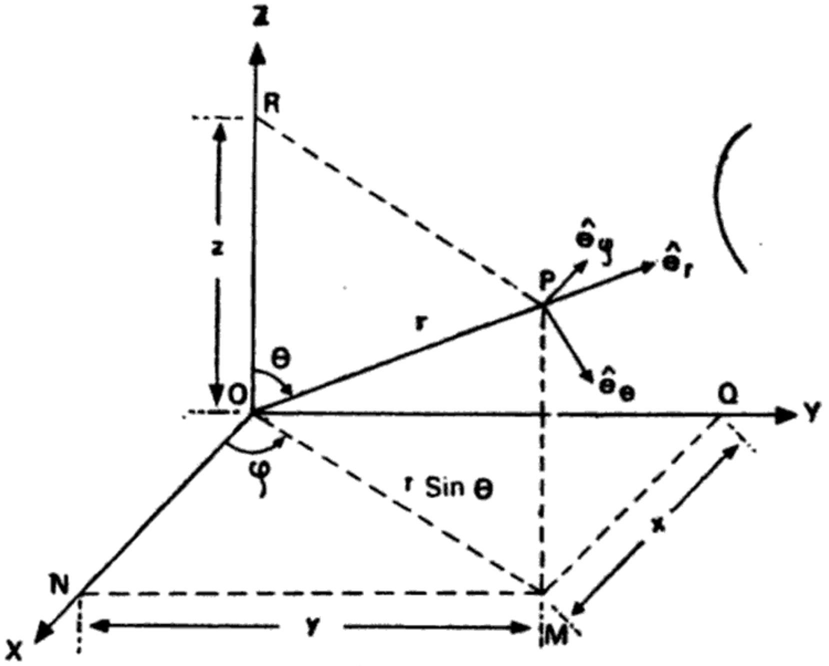


Fig. A.6 Spherical polar coordinates

Unit vectors $\hat{\mathbf{e}}_r$, $\hat{\mathbf{e}}_\theta$ and $\hat{\mathbf{e}}_\varphi$ are defined by equations

$$\hat{\mathbf{e}}_r = \frac{\partial \mathbf{r}/\partial r}{|\partial \mathbf{r}/\partial r|} = \sin\theta\cos\varphi\ \mathbf{i} + \sin\theta\sin\varphi\ \mathbf{j} + \cos\theta\ \mathbf{k} \tag{A.36}$$

$$\hat{\mathbf{e}}_\theta = \frac{\partial \mathbf{r}/\partial \theta}{|\partial \mathbf{r}/\partial \theta|} = \cos\theta\cos\varphi\ \mathbf{i} + \cos\theta\sin\varphi\ \mathbf{j} - \sin\theta\ \mathbf{k} \tag{A.37}$$

and

$$\hat{\mathbf{e}}_\varphi = \frac{\partial \mathbf{r}/\partial \varphi}{|\partial \mathbf{r}/\partial \varphi|} = -\sin\varphi\ \mathbf{i} + \cos\varphi\ \mathbf{j} \tag{A.38}$$

Thus, $\hat{\mathbf{e}}_r$, $\hat{\mathbf{e}}_\theta$ and $\hat{\mathbf{e}}_\varphi$ are the respective unit vectors pointing in the directions of the change of position vector $\mathbf{r}$ when coordinates r, θ and φ are changed by infinitesimally small amounts. These directions are shown in Fig. A.6. It can also be shown that $\hat{\mathbf{e}}_r \cdot \hat{\mathbf{e}}_\theta = 0$, $\hat{\mathbf{e}}_\theta \cdot \hat{\mathbf{e}}_\varphi = 0$ and $\hat{\mathbf{e}}_\varphi \cdot \hat{\mathbf{e}}_r = 0$. Thus, the three unit vectors are mutually perpendicular and define a right-handed coordinate system. These unit vectors satisfy the relations

$$\hat{\mathbf{e}}_r \times \hat{\mathbf{e}}_\theta = \hat{\mathbf{e}}_\varphi,\ \hat{\mathbf{e}}_\theta \times \hat{\mathbf{e}}_\varphi = \hat{\mathbf{e}}_r \text{ and } \hat{\mathbf{e}}_\varphi \times \hat{\mathbf{e}}_r = \hat{\mathbf{e}}_\theta \tag{A.39}$$

Unit vectors $\hat{\mathbf{e}}_r$, $\hat{\mathbf{e}}_\theta$ and $\hat{\mathbf{e}}_\varphi$ are not constant vectors and change in direction as θ and φ change. The time derivatives of the unit vectors can be found out to be

$$\dot{\hat{\mathbf{e}}}_r = \dot{\theta}\hat{\mathbf{e}}_\theta + \sin\theta\ \dot{\varphi}\hat{\mathbf{e}}_\varphi \tag{A.40}$$

$$\dot{\hat{\mathbf{e}}}_\theta = -\dot{\theta}\hat{\mathbf{e}}_r + \cos\theta\ \dot{\varphi}\hat{\mathbf{e}}_\varphi \tag{A.41}$$

and
$$\dot{\hat{\mathbf{e}}}_\varphi = -\sin\theta\,\dot{\varphi}\hat{\mathbf{e}}_r - \cos\theta\,\dot{\varphi}\hat{\mathbf{e}}_\theta \tag{A.42}$$

Volume element $d\tau$ is shown in Fig. A.7 and it is seen to be equal to

$$d\tau = r^2 \sin\theta\, dr\, d\theta\, d\varphi \tag{A.43}$$

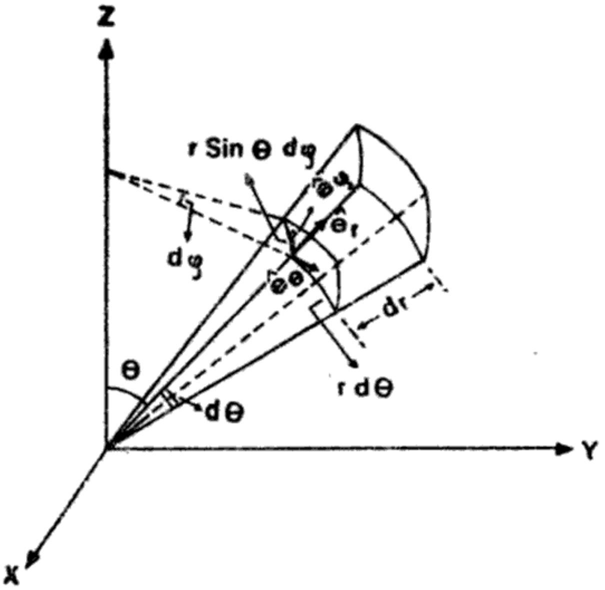


Fig. A.7 Volume element in spherical polar coordinates

The expressions for the gradient, the divergence, the Laplacian and the curl in the spherical polar coordinates are

$$\nabla\Phi = \hat{\mathbf{e}}_r \frac{\partial\Phi}{\partial r} + \hat{\mathbf{e}}_\theta \frac{\partial\Phi}{r\,\partial\theta} + \hat{\mathbf{e}}_\varphi \frac{1}{r\sin\theta}\frac{\partial\Phi}{\partial\varphi} \tag{A.44}$$

$$\nabla\cdot\mathbf{V} = \frac{1}{r^2\sin\theta}\left[\sin\theta\frac{\partial}{\partial r}(r^2 V_r) + r\frac{\partial}{\partial\theta}(\sin\theta\, V_\theta) + r\frac{\partial V_\varphi}{\partial\varphi}\right] \tag{A.45}$$

$$\nabla^2\Phi = \frac{1}{r^2\sin\theta}\left[\sin\theta\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right) + \frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Phi}{\partial\theta}\right) + \frac{1}{\sin\theta}\frac{\partial^2\Phi}{\partial\varphi^2}\right] \tag{A.46}$$

and
$$\nabla\times\mathbf{V} = \frac{1}{r^2\sin\theta}\begin{vmatrix} \hat{\mathbf{e}}_r & r\hat{\mathbf{e}}_\theta & r\sin\theta\,\hat{\mathbf{e}}_\varphi \\ \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial\theta} & \dfrac{\partial}{\partial\varphi} \\ V_r & rV_\theta & r\sin\theta\, V_\varphi \end{vmatrix} \tag{A.47}$$

A.10 VELOCITY AND ACCELERATION IN SPHERICAL POLAR COORDINATES

Position vector $\mathbf{r}$ can be represented by

$$\mathbf{r} = r\hat{\mathbf{e}}_r \tag{A.48}$$

Hence, the velocity is given by the expression

$$\begin{aligned}\mathbf{v} = \dot{\mathbf{r}} &= \dot{r}\hat{\mathbf{e}}_r + r\dot{\hat{\mathbf{e}}}_r \\ &= \dot{r}\hat{\mathbf{e}}_r + r\dot{\theta}\hat{\mathbf{e}}_\theta + r\sin\theta\,\dot{\varphi}\hat{\mathbf{e}}_\varphi\end{aligned} \tag{A.49}$$

Acceleration is given by the expression

$$\begin{aligned}
\mathbf{a} = \dot{\mathbf{v}} &= (\ddot{r}\,\hat{\mathbf{e}}_r + \dot{r}\,\dot{\hat{\mathbf{e}}}_r) + (\dot{r}\,\dot{\theta}\,\hat{\mathbf{e}}_\theta + r\ddot{\theta}\,\hat{\mathbf{e}}_\theta + r\dot{\theta}\,\dot{\hat{\mathbf{e}}}_\theta) \\
&\quad + (\dot{r}\sin\theta\,\dot{\varphi}\,\hat{\mathbf{e}}_\varphi + r\cos\theta\,\dot{\theta}\dot{\varphi}\,\hat{\mathbf{e}}_\varphi + r\sin\theta\,\ddot{\varphi}\,\hat{\mathbf{e}}_\varphi + r\sin\theta\,\dot{\varphi}\,\dot{\hat{\mathbf{e}}}_\varphi) \\
&= (\ddot{r} - r\dot{\theta}^2 - r\sin^2\theta\,\dot{\varphi}^2)\hat{\mathbf{e}}_r + (r\ddot{\theta} + 2\dot{r}\dot{\theta} - r\sin\theta\cos\theta\,\dot{\varphi}^2)\hat{\mathbf{e}}_\theta \\
&\quad + (r\sin\theta\,\ddot{\varphi} + 2\dot{r}\dot{\varphi}\sin\theta + 2r\cos\theta\,\dot{\theta}\dot{\varphi})\hat{\mathbf{e}}_\varphi
\end{aligned} \tag{A.50}$$

A.11 CIRCULAR CYLINDRICAL OR CYLINDRICAL POLAR COORDINATES

In this coordinate system (Fig. A.8), we have the three coordinates as

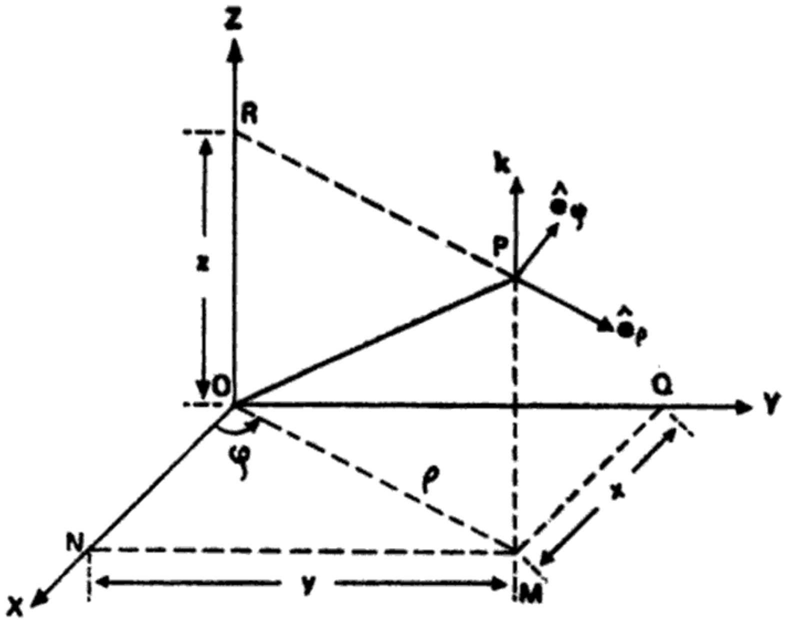


Fig. A.8 Cylindrical polar coordinates

$$q_1 = \rho, \qquad 0 \leqslant \rho \leqslant \infty$$
$$q_2 = \varphi, \qquad 0 \leqslant \varphi \leqslant 2\pi$$

and
$$q_3 = z, \quad -\infty \leqslant z \leqslant \infty$$

These coordinates are related to the cartesian coordinates by the relations

$$\left.\begin{aligned} x &= \rho\cos\varphi \\ y &= \rho\sin\varphi \\ z &= z \end{aligned}\right\} \tag{A.51}$$

It can be easily observed that

(i) $\rho = \sqrt{x^2 + y^2}$ = constant describe a family of right-circular cylinders with the axis of z as the common axis,

(ii) $\varphi = \tan^{-1}\dfrac{y}{x}$ = constant are the half-planes through the polar axis, viz. z-axis, and

(iii) z = constant are the planes perpendicular to the z-axis.

The scale factors for this coordinate system are

$$h_1 = h_\rho = 1,\ h_2 = h_\varphi = \rho \text{ and } h_3 = h_z = 1 \tag{A.52}$$

CLASSICAL MECHANICS is a lucid introduction to Vectors, Classical Mechanics and Special Theory of Relativity. It gives the student a firm grounding for advanced courses on Quantum Mechanics, Nuclear Physics, Electrodynamics, etc. Apart from covering basic topics like Vector Algebra and Analysis, Mechanics of Single and Systems of Particles, and Motion of Rigid Body, it comprehensively discusses Langrangian Formulation, Hamilton's Equations, Collisions of Particles and Theory of Small Oscillations. It contains a systematic treatment of Actions of Particle under different types of Forces and gives a unified approach to Fields and Potentials in Inverse Square Law Force Field. Included is an appendix on Coordinate System and an extensive Bibliography.

The book meets complete requirements of undergraduate and post graduate students. It contains numerous questions and problems at the end of each chapter which should provide an invigorating exercise for the student and teacher.

R G TAKWALE is currently Vice Chancellor, Indira Gandhi National Open University (IGNOU), New Delhi. Earlier he was Professor of Physics at the University of Poona where he has been teaching for over twenty years. He obtained his Ph D in Theoretical Physics from Moscow State University in 1965. He is the founder member of the Matharashtra Academy of Sciences and is member of the American Institute of Physics, Indian Science Congress Association and Indian Physics Association. He is the main author of *An Introduction to Physics* (1976) and the co-author of *A Course in Mathematics for Students of Physical Sciences (1977).* His current fields of interest include interaction of Particles and Fields with Matter, Cerenkov Radiation and Physics Education.

P S PURANIK is currently a voluntary research worker at the Department of Physics, University of Poona. He got his M Sc from the University of Poona and has taught at the Fergusson College, Poona, Saraswati Bhavan College, Auragabad and M E S College, of Arts and Sciences, Poona. His present research interests are in Solid State Physics (X-Ray Spectroscopy). He is the co-author of *Principles of Physics* (1970).

The McGraw-Hill Companies

Tata McGraw-Hill
Publishing Company Limited
7 West Patel Nagar, New Delhi 110 008

Visit our website at : www.tatamcgrawhill.com

ISBN 0-07-096617-6

9780070966178